

AFIT/GOR/ENS/98M-19

VERIFICATION AND VALIDATION OF
FAARR MODEL AND DATA ENVELOPMENT
ANALYSIS MODELS FOR UNITED STATES
ARMY RECRUITING

THESIS

Gene M. Piskator
Captain, U.S. Army

AFIT/GOR/ENS/98M-19

Approved for public release; distribution unlimited

DTIC QUALITY INSPECTED 2

19980427 122

The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government.

THESIS APPROVAL

STUDENT: Gene M. Piskator, U.S. Army

CLASS: GOR-98M

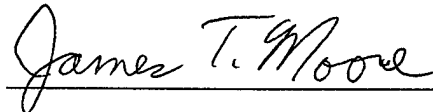
THESIS TITLE: Verification and Validation of FAARR Model and Data Envelopment
Analysis Models for United States Army Recruiting

DEFENSE DATE: 20 February 1998

Committee: Name/Title/Department

Signature

Co-Advisor James T. Moore
Associate Professor
Department of Operational Sciences

Handwritten signature of James T. Moore in cursive script, underlined.

Co-advisor Jack M. Kloeber Jr.
Assistant Professor
Department of Operational Sciences

Handwritten signature of Jack M. Kloeber Jr. in cursive script, underlined.

AFIT/GOR/ENS/98M-19

VERIFICATION AND VALIDATION OF FAARR MODEL AND DATA
ENVELOPMENT ANALYSIS MODELS FOR UNITED STATES ARMY
RECRUITING

THESIS

Presented to the Faculty of the Graduate School of Engineering
of the Air Force Institute of Technology

Air University

In Partial Fulfillment of the Requirements for the
Degree of Master of Science in Operations Research

Gene M. Piskator, B.S.

Captain, US Army

March 1998

Approved for public release; distribution unlimited

Preface

The purpose of this research was to determine the accuracy of the forecast estimates from the Army's Forecast and Allocation of Army Recruiting Resources (FAARR) decision support system and to develop a Data Envelopment Analysis (DEA) modeling strategy which produces accurate DEA models. By doing so, I hoped to identify the most accurate DEA model formulation to estimate recruiting battalion efficiency as well as to illustrate the use of this DEA efficiency information in econometric forecasting models.

I could not have conducted this research without the assistance and support of others. I would like to thank my co-advisors, LtCol James T. Moore and LTC Jack M. Kloeber Jr., for their astute direction, professional guidance, and frank criticism. By providing me the benefit of their years of analytical experience and allowing me the flexibility to examine all aspects of DEA, this research is a better product. I would also like to thank LTC Gregory Hoscheit of the United States Army Recruiting Command for his assistance and recommendations. Without his technical insights, knowledge of the recruiting process, or timely response to my requests, this research could not have been completed. Finally, I wish to thank my father, Walter Piskator, who taught me the value of hard work and an appreciation for education, and my wife, Tara, for her love, her encouragement, and her ability to make me see the important aspects of life.

Gene M. Piskator

Table of Contents

	Page
Preface.....	ii
Table of Contents.....	iii
List of Figures.....	vi
List of Tables.....	vii
Abstract.....	ix
I. Introduction.....	1
1.1 Background.....	1
1.2 Research Importance.....	3
1.3 Research Objectives.....	5
1.4 Research Questions.....	6
1.5 Research Scope.....	7
1.6 Assumptions.....	7
1.7 Overview and Format.....	8
1.8 Research History.....	9
II. Literature Review.....	10
2.1 Production FunctionDefinition.....	10
2.2 Data Envelopment Analysis Basics.....	12
2.3 Deterministic versus Stochastic DEA Models.....	28
2.4 DEA Sensitivity Analysis.....	31

2.5 Beyond the Basic DEA Model.....	35
2.6 FAARR Model Background.....	39
III. Methodology.....	47
3.1 Introduction.....	47
3.2 Input Data Analysis.....	48
3.3 DEA Model Linear Programming Constraints and Resource Level Sensitivity Analysis.....	50
3.4 FAARR Model Validation Forecasts of Actual Contract Production.....	57
3.5 FAARR Model Estimation of Production Function Parameters and DEA Efficiency Scores from the Simulation of a Known Production Function.....	60
3.6 FAARR Model Validation Summary.....	66
3.7 A Method for Selecting an Accurate DEA Model Formulation.....	69
IV. Results of the DEA Modeling Strategy.....	78
4.1 Identifying Relevant DEA Input Variables.....	78
4.2 Estimating a Recruiting Production Function.....	82
4.3 Simulation of Recruiting Production Function.....	86
4.4 Simulation Results.....	89
4.5 Efficiency Estimates Using CCR Model Formulation.....	94
4.6 Practical Application of DEA Efficiency Information in Combined OLS/DEA Model.....	96
V. Conclusions and Recommendations.....	102
5.1 Conclusion.....	102
5.2 Improving USAREC Econometric and Forecasting Models.....	105
5.3 Extensions of Current Research.....	106

Appendix A. Sample GAMS Simulation Model.....	109
Bibliography.....	115

List of Figures

Figure	Page
2.1 Data Envelopment Analysis Empirical Efficiency Frontier.....	13
2.2 DEA Efficient Frontier Model vs. Regression Model.....	15
2.3 Technical versus Allocative Efficiency.....	17
2.4 DEA Additive Model Envelopment Surface.....	18
2.5 DEA Multiplicative Model Envelopment Surface.....	19
2.6 DEA Input Oriented CCR Model Envelopment Surface.....	21
2.7 DEA Output Oriented CCR Model Envelopment Surface.....	21
2.8 DEA Input Oriented BCC Model Envelopment Surface.....	23
2.9 DEA Output Oriented BCC Model Envelopment Surface.....	23
2.10 DEA Super-Efficiency Model Envelopment Surface.....	25
2.11 Army Forecast and Allocation of Recruiting Resources (FAARR) Model.....	40
3.1 FAARR Model Resource Level Affect on Contract Production Forecast.....	56
3.2 FAARR Forecast Evaluation of Simulated Function.....	66
4.1 Recruiting Data PCA Eigenvalue Scree Test.....	79
4.2 Actual GSMA Contract Production versus OLS Estimate.....	84
4.3 Actual GSMA Contract Production versus Efficient Frontier Benchmarking Estimate.....	86

List of Tables

Table	Page
3.1 Recruiting Resource DEA Virtual Multiplier Bounds.....	52
3.2 GSMA Contract Forecast Sensitivity Analysis.....	54
3.3 GSMA Contract Forecast Percentage Change.....	54
3.4 FAARR Model Forecast MOEs using DEA Virtual Multipliers from Same Quarter in Previous Year.....	58
3.5 FAARR Model Forecast MOEs using DEA Virtual Multipliers from Previous Quarter.....	59
3.6 FAARR Model Forecast MOEs using Average DEA Virtual Multipliers for all Four Quarters from Previous Year.....	59
3.7 Recruiting Battalion Naive Forecast 1 Results.....	60
3.8 Simulation Model Results Using FAARR Model to Estimate Efficiency Scores and Parameters from a Known Production Function.....	63
3.9 Simulation Model 1 Confusion Matrix.....	65
3.10 Simulation Model 2 Confusion Matrix.....	65
4.1 Recruiting Data PCA Factor Loadings Matrix.....	78
4.2 OLS Statistical Significance Results for Recruiting Resource Variables.....	79
4.3 Recruiting Data Correlation Matrix.....	80
4.4 Recruiting Resource Data Variance Inflation Factors.....	81
4.5 Recruiting Resource Variable Analysis and Screening.....	81
4.6 Simulation Results for OLS Production Function without Error Term.....	89
4.7 Simulation Results for Frontier Production Function without Error Term.....	89

4.8 Simulation Results for OLS Production Function without Error Term.....	90
4.9 Simulation Results for Frontier Production Function with Error Term.....	90
4.10 Average Simulation Results For All Production Functions.....	90
4.11 Simulation Result Summary.....	92
4.12 Comparison of DEA Model Efficiency Scores.....	94
4.13 Comparison of Average Forecast Results for OLS and OLS/DEA Models.....	99

Abstract

This research has two objectives—to verify and validate the U.S. Army's Forecast and Allocation of Army Recruiting Resources (FAARR) model and to develop a Data Envelopment Analysis (DEA) modeling strategy.

First, the FAARR model was verified using a simulation of a known production function and validated using sensitivity analysis and *ex-post* forecasts. FAARR model forecasts were not accurate and were extremely sensitive to any changes in the model's linear programming constraints and to changes in recruiting resource levels.

Second, this research describes a three phase modeling strategy to build accurate DEA models. DEA has become a popular tool to evaluate the relative efficiency of many types of organizations. However, the literature has paid little attention to the practical problems of selecting the appropriate input variables and envelopment frontier. Analysts may use a number of diagnostic techniques to detect mis-specification in statistics based models. No such diagnostics exist for DEA models. Without *a-priori* knowledge concerning the production process's appropriate input variables and returns to scale, analysts do not know if they have constructed an accurate DEA model. Using a three-phase strategy, relevant DEA model input variables are selected using Principal Component Analysis and Ordinary Least Squares (OLS) regression. The appropriate DEA envelopment frontier is selected using a Monte-Carlo simulation of an estimated production function representing the actual production process. The research concludes by demonstrating *ex-post* forecasts from a combined OLS/DEA model were more accurate when the DEA model formulation selected by the three phase modeling strategy was used.

VERIFICATION AND VALIDATION OF FAARR MODEL AND DATA
ENVELOPMENT ANALYSIS MODELS FOR UNITED STATES ARMY
RECRUITING

I. Introduction:

“The fact is, there are only two qualities in the world: efficiency and inefficiency, and only two sorts of people: the efficient and the inefficient.”

George Bernard Shaw, *John Bull's Other Island*, 1907.

1.1 Background

More than any other organization in the Department of Defense, the United States Army relies on a large annual cohort of new enlistees in order to maintain a viable fighting force. Of all the recruits entering active military service in any year, 45% will join the Army (39:16). Including the Reserve Components, the Army recruits more personnel each year than all other Department of Defense services combined (22). In Fiscal Year (FY) 1997, the Army's recruiting mission for the Active and Reserve components was almost 139,000 soldiers (51). The Army's difficult, unglamorous, and sometimes dangerous mission makes recruiting quality personnel a challenge.

Since its transition to an all volunteer force in 1974, the Army has had a difficult mission of attracting this large cohort of high quality new recruits (53:568). General economic prosperity, potential world-wide conflicts, American youths' attitudes toward the military (22), a decreasing youth population, and a “rightsizing” military establishment all combine to make the Army's recruiting mission more difficult.

Although tremendously beneficial to the populous in general, continuing domestic economic prosperity impedes Army recruiting. December 1997 Labor Department reports indicate the nation's unemployment rate had fallen to 4.6 percent, its lowest level since December 1973 (45:a2). Army marketing research suggests for each ten percent drop in the number of unemployed adults, there is a corresponding seven to nine percent drop in the number of people interested in joining the military (3:22).

Continued technological advances in ground warfare and evolutionary changes in fighting doctrine require a higher quality, more intelligent soldier--one who is capable of operating and maintaining sophisticated weapons systems and skilled in the use of computers. In order to meet its requirement for high quality new enlistees, the Army's goal is to recruit individuals who score in the top fiftieth percentile for intelligence on the Armed Forces Qualification Test (AFQT) (50).

Due to these market forces and the Army's rising intelligence standards, the resources which the Army has committed to recruiting-- both in terms of dollars and trained personnel-- have increased. Although the Army as a whole is continuing to downsize and military budgets continue to decline, the Army has been forced to increase the resources committed to recruiting in order to achieve its recruiting goals. For FY97, United States Army Recruiting Command's (USAREC) advertising budget increased \$15 million, to \$86 million, and the total number of recruiters increased from 5200 to 5225 (48:15). The Army recently contracted with its advertising agency, Young and Rubicam, for \$440 million of advertising services through FY02 (49:22). Recruiting young people into military service is a major industry and requires significant amounts of the Department of

Defense's resources. Each new recruit costs the Department of Defense approximately \$6000 for recruiters, advertising, education benefits and bonuses (22). Using this information in a conservative estimate, the United States Army will commit over \$4.0 Billion in resources to recruit new soldiers between FY98 and FY02.

1.2 Research Importance

As military personnel and budgetary resources continue to decline, it is increasingly important for the Army to efficiently utilize its recruiting resources to enlist a sufficient number of the highest quality recruits. In order to estimate marginal returns to production (elasticities) for additional resources, analysts usually rely on economic models to estimate these parameters for each specific resource (35:208). Commonly referred to as causal models, regression based econometric models may also be used to forecast future production based upon minor changes in resource levels (37:185). Time series forecasting models-- either smoothing methods or Box-Jenkins autoregressive models --are less computationally complex models which rely on the past behavior of the variable being predicted to estimate forecasts (35:205-206). These types of models do not provide parameter estimates, but they may provide more accurate short term forecasts than causal models (35:210). Usually, the type of model developed-- either causal or time-series-- is based upon its primary intended purpose-- parameter estimation or forecasting (35:208-210). In order to effectively allocate its limited recruiting resources, USAREC requires accurate information on both the resource parameters and contract forecasts.

The Center for Cybernetic Studies at the University of Texas at Austin developed the Forecast and Allocation of Army Recruiting Resources (FAARR) decision support system for the Army's Recruiting Command to provide information on both forecasts and resource allocation. This model is a Personal Computer (PC) platform based, deterministic, two-stage, Data Envelopment Analysis (DEA) and linear optimization system which forecasts either contract output or resource requirements (13:9). USAREC leadership asked the Air Force Institute of Technology's Operational Sciences Department to evaluate the robustness of this model and measure the model's forecast accuracy prior to USAREC using it to assist decision makers.

The Operations Research community has found many new uses of DEA efficiency information in multiple stage mathematical models. For example, researchers have combined DEA results with goal programming (29) and regression techniques (2). However, the literature has paid minimal attention to specific procedures or modeling strategies to build accurate DEA models (46:233). There are few formal procedures or heuristics to select both the appropriate input variables and the shape of the envelopment frontier to develop an accurate DEA model. Analysts may use a number of diagnostic techniques to detect misspecification in statistics based models including analysis of residuals, adjusted R^2 , and the C_p criteria, to name a few (23:235). No such diagnostics exist for DEA models. Without *a-priori* knowledge of the production process' appropriate input resources and returns-to-scale classification, analysts do not know if they have constructed an accurate DEA model. If DEA efficiency information is to be useful as a management tool or as an input for multiple stage mathematical models,

analysts need to be confident that the specific DEA model accurately classifies the evaluated entities as efficient or inefficient.

1.3 *Research Objectives*

The purpose of this research is fourfold:

1. Verify and validate the FAARR decision support system and determine the accuracy and robustness of the model's forecast estimates.
2. Develop a Data Envelopment Analysis modeling strategy which produces more accurate DEA models.
3. From a set of alternate models, identify the most accurate DEA model formulation to estimate recruiting battalion efficiency.
4. Illustrate the use of DEA efficiency information in causal OLS forecasting models.

This research describes a three step statistical and simulation based strategy to develop an accurate and robust DEA model. The DEA model may then be used to identify efficient and inefficient recruiting battalions. This information may be used with other, second stage mathematical models to more accurately estimate input resource elasticities, forecast contract production, or allow the optimal reallocation of recruiting resources among recruiting battalions.

The following steps will be used to achieve the first objective--verification and validation of the FAARR model:

1. Analyze the recruiting resource and contract production data sets.

2. Conduct sensitivity analysis to determine the change in forecasted production due to changes in the DEA model virtual multiplier constraints and changes in aggregate recruiting resource levels.
3. Use past resource and production data with the FAARR model to evaluate the model's accuracy.
4. Use a Monte-Carlo simulation of a known production function to determine the ability of the FAARR model to accurately estimate production function parameters.

The following three stage strategy was developed to identify the most accurate and robust DEA model formulation to be used in the efficiency assessment of U.S. Army recruiting battalions:

1. The use of Principal Component Analysis and Ordinary Least Squares (OLS) regression to screen and select appropriate DEA input variables.
2. Monte-Carlo simulation of a production function which approximates the production process to select the most accurate DEA envelopment surface.
3. Use of DEA derived efficiency information in an illustrative Ordinary Least Squares model to demonstrate model improvement as measured by increased forecast accuracy.

1.4 *Research Questions*

- Does the FAARR model accurately predict contract production?
- How sensitive is the model to changes in quarter to quarter resource allocation and production function parameter assumptions?

- How sensitive is the FAARR estimated contract forecasts to changes in the aggregate recruiting resource levels?
- Which DEA model formulation most accurately estimates actual recruiting battalion efficiency?
- How can an analyst select an accurate DEA model formulation given a selection of input and output variables and various envelopment frontiers?

1.5 *Research Scope*

An attempt to validate the accuracy and applicability of the FAARR model was conducted using a data set for USAREC recruiting battalions consisting of quarterly data from 1st Quarter FY96 thru 3rd Quarter FY97. The OLS model developed in this research is for illustrative and comparative purposes only, and is not intended to represent the most accurate GSMA forecasting model available.

1.6 *Assumptions*

The following assumptions were necessitated during the research process:

1. The U.S. Army recruiting process can be modeled as a production process using a mathematical production function.
2. Recruiting battalion leaders and recruiters attempt to maximize quarterly enlistment contract production given any allocation of recruiting resources.
3. Historical USAREC supplied input and production data is accurate and deterministic in nature.

1.7 *Overview and Format*

This research is presented in the following manner:

Chapter 1 introduces the research, provides recruiting environment background information, lists the research assumptions, defines the scope, and presents the assumptions.

Chapter 2 presents the literature review and includes a definition of production functions; a review of DEA theory; DEA assumptions, advantages, and limitations; and classical DEA model formulations. Additionally, Chapter 2 discusses stochastic DEA, DEA sensitivity analysis, and use of DEA efficiency information in multiple stage mathematical models. Finally, Chapter 2 outlines the FAARR model, its assumptions, and mathematical formulations.

Chapter 3 describes the FAARR model validation and verification process and outlines a three stage DEA model building methodology. This strategy includes the use of statistical analysis, production function estimation, and Monte-Carlo simulation.

Chapter 4 illustrates the use of the three stage DEA model building strategy for the Army recruiting process. Chapter 4 also describes the specific model building steps, variable selection logic, results of the DEA model, and use of the DEA efficiency information in a causal OLS model.

Chapter 5 summarizes both the results of the FAARR model verification and validation analysis as well as the results of the DEA model building strategy. Finally, Chapter 5 suggests future research to analyze the selection of an accurate DEA model.

1.8 *Research History*

As with much research, the direction and scope of this particular research changed throughout the research process. When this research began in April 1997, its objective was to estimate a confidence interval for the FAARR model contract forecast. Since the FAARR model is a deterministic model, it was thought that boot-strapping or Monte-Carlo simulation may be an appropriate solution methodology for estimating the confidence interval. However, as the research progressed, the accuracy of the FAARR model's forecasts and the validity of the model's assumptions were brought into question. It became obvious that an accurate estimate of the confidence interval for a biased, inaccurate forecast would not be useful to the USAREC. After concluding the FAARR model was not valid, the research focus shifted to identifying which DEA model most accurately estimated efficiency of U.S. Army recruiting battalions. This document presents results from all phases of the analysis to include FAARR model verification and validation, identifying accurate DEA models, and use of DEA efficiency information.

II. Literature Review

2.1 Production Function Definition

A production function describes the functional relationship between a production processes' inputs and outputs. Usually, a production function is defined as a schedule, table, or mathematical function which catalogs the efficient output possibilities for the production process (38:173).

Production functions can be formulated as non-parametric or parametric mathematical functions. DEA models belong to the class of non-parametric production functions. The functional relationship between resource input and production output need only be monotonic and concave (47:104). Parametric production functions assume a more specific and restrictive functional form and can be formulated as linear functions, log-linear functions, or log-log (Cobb-Douglas) functions.

Analysts can use a number of different models to estimate production functions. These models include, but are not limited to, statistical methods using Ordinary Least Squares (OLS) regression, linear programming methods using an efficient frontier benchmarking formulation, or more advanced mathematical programming methods to estimate stochastic frontiers.

Production functions are widely categorized based on their returns-to-scale properties. Returns-to-scale is an economic term which defines the general ability of the production process to convert inputs to outputs. Returns-to-scale are generally Increasing (IRS), Decreasing (DRS) or Constant (CRS). For a production process which exhibits IRS, any

percentage increase in inputs results in a greater percentage increase in outputs. For example, for an IRS process, if the manufacturer increases all resource inputs by 5%, overall output will increase by more than 5%. For a CRS process, any percentage increase in resources will result in a similar percentage increase in outputs (38: 207-208). A Variable Returns to Scale (VRS) production process changes its returns to scale property at varying levels of production (16:71).

This research assumes we can use a mathematical production function in some to be determined form to model the U.S. Army recruiting process. One commonly used mathematical production function form is Cobb-Douglas. Cobb-Douglas functional forms possess many desirable qualities and are extensively used by the operations research and economic communities (25:299). Cobb-Douglas forms assume constant elasticity of substitution between resource inputs (18:3748) and can be used to readily compute input and output elasticities. Computationally, OLS or linear programming may be used to estimate Cobb-Douglas functions by simply using the natural logarithm of the applicable variables (35:73). Cobb-Douglas functional forms have been used to empirically analyze production and distribution economics (18:3747) and educational programs (2:259).

Cobb-Douglas production functions usually take the form: $y = \alpha_0 \prod x_i^{\beta_i}$, with x_i representing resource inputs and y representing produced outputs. Ratios of the estimated coefficients (β_i 's)-- which represent resource output elasticities (18:3747)-- can be used to calculate Marginal Rates of Substitution (MRS) (6:34) between resource inputs.

Additionally, the sum of the resource output elasticities indicate whether the functions is IRS, CRS, or DRS. A Cobb-Douglas functional form is IRS if the sum of the estimated

coefficients (β_i) is greater than 1. The functional form is CRS if the sum of the estimated coefficients is equal to 1 and the functional form is DRS if the sum of the estimated coefficients is less than 1 (38: 207-208).

2.2 Data Envelopment Analysis Basics

Data Envelopment Analysis (DEA) is a descriptive mathematical modeling methodology that determines a Decision Making Unit's (DMU) efficiency using linear programming techniques. Decision Making Units are comparable productive entities within an organization which transform the same measurable inputs into measurable outputs. DMUs operate in a similar environment and DMUs' management decisions are guided by similar measurable objectives (29:171). Bank branches, warehouses, schools, or Army recruiting battalions are examples of DMUs.

Originally developed by Charnes, Cooper, and Rhodes in 1978, DEA models calculate an empirical non-parametric production frontier by comparing each DMU's resource inputs and produced outputs (47:104). DEA models are loosely based upon classical production theory (15:44). The theoretical constructs of the resource input/production output ratio and production possibility frontier of DEA date back to the work on technical efficiency by the economist, M. J. Farrell, in 1957 (26).

The DEA model determines a relative efficiency rating for each DMU by calculating an efficiency score which represents the difference between a specific DMU's outputs and resource inputs compared to the inputs and outputs observed among all other DMUs. An efficient DMU produces the maximum observed output given its resource inputs and has

an efficiency rating of 1. In essence, DEA efficiency is no more than Pareto optimality. A DMU can not be efficient if there is another DMU-- or virtual DMU-- which produces the same amount of output with less resource input (44:6). Figure 2.1 illustrates the empirical production frontier (a BCC envelopment) and the efficient and inefficient DMUs using the Pareto optimality efficiency criteria. For example, DMU B produces 5 units of output using 3 units of input (resources). In contrast, DMU E produces 3 units of output using 5 units of input. DMU B can actually produce more output with less input. Therefore, DMU E is inefficient.

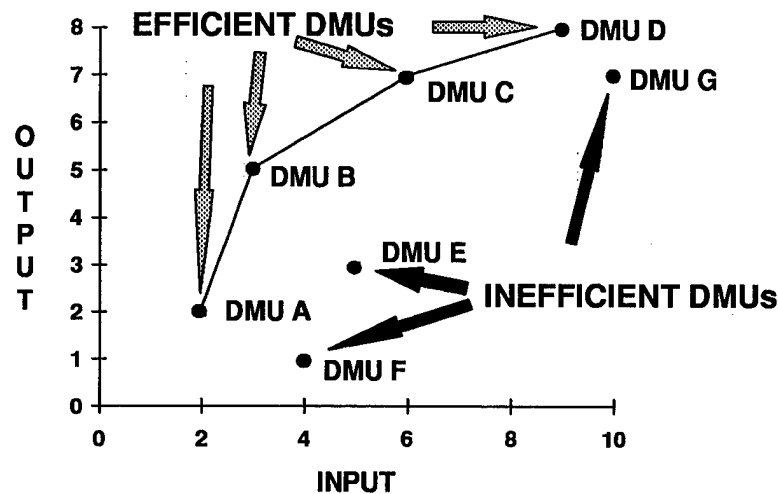


Figure 2.1: Data Envelopment Analysis Empirical Efficiency Frontier

DEA models are classified as non-parametric models and place minimal assumptions on the DMU's "theoretical" underlying production function. Unlike classical econometric techniques which stipulate a specific, theoretical functional form for the production function --usually Cobb-Douglas or Constant Elasticity of Substitution (CES)--DEA methods do not specify a functional form or a specific distribution of an error term

(47:104). An individual DMU's "production function", or relationship between inputs and outputs, needs to be monotonic and concave.

In order to calculate the efficiency of a particular DMU, analysts use linear programming to determine virtual multipliers or "weights" for the relative value of the various outputs and inputs that maximize a specific DMU's efficiency score. The DMU's efficiency score measures the distance a DMU lies from the efficient frontier (16:26). A DMU with an efficiency score of 1 lies on-- and therefore determines-- the efficient frontier. A particular DMU, say DMU_1 , may "choose" any combination of input and output weights (virtual multipliers) in order to maximize its own efficiency score subject to the constraint that all other DMUs' efficiency ratings using DMU_1 's particular weights and the other DMU's resource inputs are feasible (7:247). A separate linear programming formulation is used to calculate the efficiency score for each DMU. DEA efficiency estimates are calculated from observed data for each DMU and produce only relative efficiency measures in comparison to all other DMUs.

The efficient DMUs form an envelopment surface or production possibility frontier (42:442). The efficient production frontier is not a theoretical efficient frontier, but an unambiguous relative frontier calculated from the actual observed performance (output) of some subset of the DMUs being evaluated. Unlike classical regression techniques that estimate an average production function across an entire industry, DEA techniques identify two mutually exclusive subsets of DMUs--efficient and inefficient. The efficient DMUs "map out" or determine the relative efficient production frontier. As Stolf states, "...DEA is a methodology directed to frontiers rather than central tendencies." (47:108)

Figure 2.2 illustrates the difference between DEA and regression based estimates of the production function.

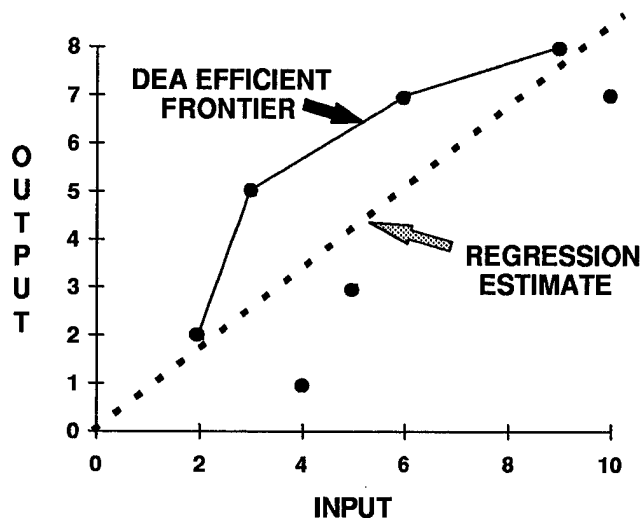


Figure 2.2: DEA Efficient Frontier Model vs. Regression Model

One common misperception concerning DEA is that the technique accurately identifies both efficient and inefficient units. This is not the case. DEA identifies inefficient units and may identify efficient units. As Golany and Yu state (29:179):

If DEA identifies a DMU as inefficient, it means it has found evidence (i.e., other efficient DMUs) to its inferior position. On the other hand, if DEA identifies a DMU as efficient it only means that it is unable to find evidence in the observed data, [that the DMU is inefficient] but it does not imply that this DMU is indeed efficient with respect to the unknown production function.

Using an analogy to the American judicial system, a DMU is considered to be efficient until proven inefficient.

A DMU may be technically efficient--have a DEA score of 1--due to an unrealistic selection of resource weights. As a result, certain DMUs may be rated as efficient solely due to a single input or output, even though that input or output may be relatively

unimportant. Although technically efficient, the DMU is actually allocatively inefficient (16:205). Figure 2.3 illustrates the concepts of technical and allocative efficiency. All Pareto optimal DMUs-- identified as technically efficient-- have an efficiency score of 1. However, not all technically efficient DMUs are necessarily allocatively efficient. For example, say we know the "market value"-- marginal revenue of the output or consumer cost-- of an additional unit of output is 1.5 units of resource input. This is depicted by the dotted Marginal Revenue of Output line in Figure 2.3. As a manufacturer, we would not use large DMUs-- DMUs which require more than a total of 6 units of resource input-- to produce more output. As DMUs inputs are increased, the production process exhibits decreasing returns-to-scale. The additional revenue for any additional unit of output is actually less than what it would cost us to produce that output. For example, DMU D is producing 8 units of output using 9 units of input and is technically efficient--no other unit can produce more output with less input. However, the marginal revenue of that extra unit of output-- 1.5 units of resource input-- is not worth its marginal cost-- 3 units of resource input. Therefore, because it costs more to produce the last unit of output than what that last unit is actually worth, DMU D is allocatively inefficient. The problem identifying allocatively inefficient DMUs is that the analyst is rarely able to specify an actual cost in output for each resource input for non-profit oriented organizations. Allocative efficiency requires knowledge of how to "cost out" the various input and outputs.

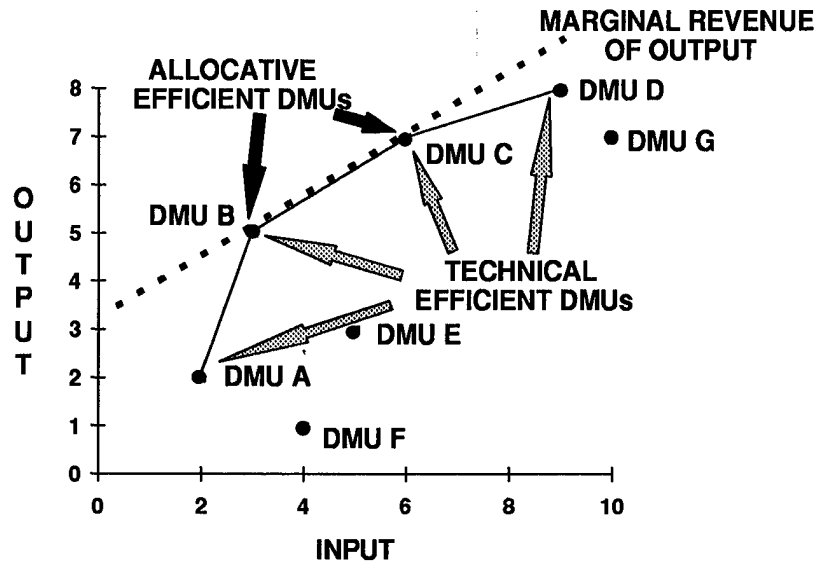


Figure 2.3: Technical versus Allocative Efficiency

Depending upon the analyst's assumptions concerning the industry wide returns-to-scale, the geometry of the efficient envelopment surface, and the projection of inefficient DMUs onto the efficient frontier, there are several variations of the basic DEA model. These models determine production frontiers with different shapes (16:45) and may produce radically different DMU efficiency scores.

The most common formulations of DEA models are the Additive, Multiplicative, CCR, and BCC models. The optimal value from the solution of the Additive (1985) DEA model formulation calculates an efficiency rating that measures the rectilinear distance a particular DMU lies from the closest DMU on the efficient frontier. The efficient DMU must produce at least as much output as the inefficient DMU. In other words, the efficient DMU lies in a "Northwesterly" direction compared to the inefficient DMU (16:28). The Additive model produces a piece-wise linear production frontier and has Variable Returns-to-Scale (16:28) as depicted in Figure 2.4.

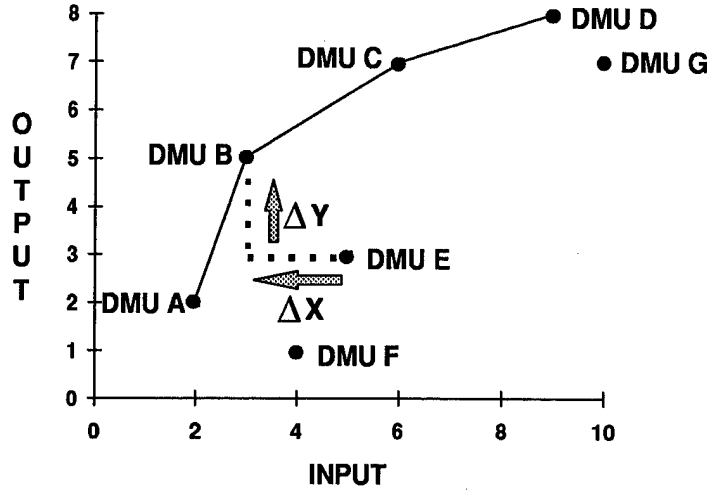


Figure 2.4: DEA Additive Model Envelopment Surface

The dual mathematical formulation of the Additive DEA model is:

$$\begin{aligned}
 &\text{Maximize: } z = \sum_r y_{rk} u_r - \sum_i x_{ik} v_i + u_o \\
 &\quad \text{subject to} \\
 &\sum_r y_{rj} u_r - \sum_i x_{ij} v_i + u_o \leq 0, \text{ for all } j=1, \dots, n \\
 &\quad u_r \geq 1, \text{ for all } r \\
 &\quad v_i \geq 1, \text{ for all } i
 \end{aligned}$$

with¹

- $z \equiv$ efficiency score of DMU k
- $y_{rk} \equiv$ output r for DMU k
- $x_{ik} \equiv$ input i for DMU k
- $u_r \equiv$ virtual multiplier for output r
- $v_k \equiv$ virtual multiplier for input i
- $u_o \equiv$ free intercept term for DMU k
- $n \equiv$ total number of DMUs being evaluated

¹ The definitions and notation of the terms used in this DEA math formulation are generally standard across the DEA literature. This notation is used throughout the thesis. Additionally, the author uses the Dual linear programming formulation because of its intuitive economic similarity to the Cobb-Douglas production function.

The Multiplicative DEA model produces a piecewise log-linear production possibility frontier. This model is similar to the Additive model, but all inputs and outputs are expressed as the natural logarithms of the original data. The basic Multiplicative DEA model may be specified with Constant Returns-to-Scale (CRS) by not using an intercept term in the mathematical formulation, or the model may be specified with Variable Returns-to-Scale (VRS) using an intercept term in the mathematical formulation (16:30). The model formulation results in a functional form which is analogous to a Cobb-Douglas production function found in classical economic theory (20:529). Figure 2.5 illustrates the shape of the empirical possibility frontier.

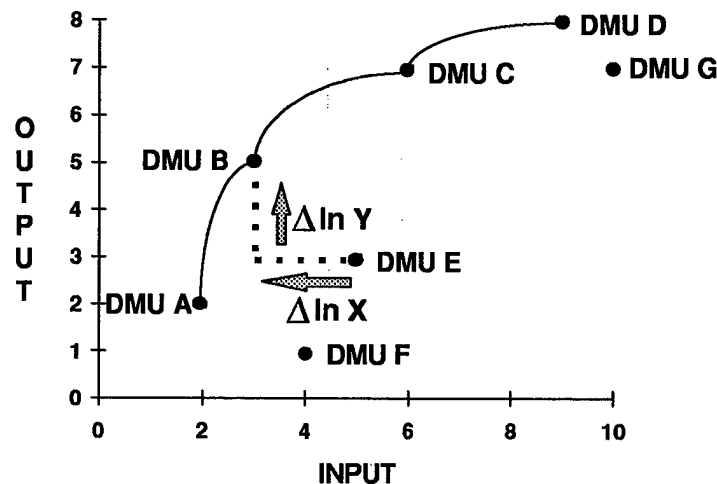


Figure 2.5: DEA Multiplicative Model Envelopment Surface

The mathematical formulation of the Multiplicative DEA model (VRS) is:

$$\text{Maximize: } z = \sum_r \ln(y_{rk})u_r - \sum_i \ln(x_{ik})v_i + u_o$$

subject to

$$\sum_r \ln(y_{rj})u_r - \sum_i \ln(x_{ij})v_i + u_o \leq 0, \text{ for all } j=1, \dots, n$$

$$u_r \geq 1, \text{ for all } r$$

$$v_i \geq 1, \text{ for all } i$$

Likewise, the mathematical formulation of the Multiplicative model without an intercept term (CRS) is:

$$\text{Maximize: } z = \sum_r \ln(y_{rk})u_r - \sum_i \ln(x_{ik})v_i$$

subject to

$$\sum_r \ln(y_{rj})u_r - \sum_i \ln(x_{ij})v_i \leq 0, \text{ for all } j=1, \dots, n$$

$$u_r \geq 1, \text{ for all } r$$

$$v_i \geq 1, \text{ for all } i$$

The Charnes, Cooper and Rhodes (CCR) DEA model (1978) results in a linear, constant returns-to-scale envelopment surface. The CCR model formulation can be either input oriented or output oriented. The two forms provide different projections of inefficient DMUs onto the empirical efficient frontier. The specific form chosen depends upon how management intends to use the efficiency information. The input-orientation focuses on maximal movement toward the efficiency frontier through proportional reduction of inputs and the output-orientation focuses on maximal movement toward the efficiency frontier by proportional augmentation of outputs (16:37). The efficiency scores from the CCR model measure the distance to a point on the efficient frontier. This point may represent an actual DMU or a virtual DMU. The CCR model assumes efficient production is theoretically possible at any point along the efficient frontier. A graphical

representation of the input and output oriented CCR models are depicted in Figure 2.6 and Figure 2.7.

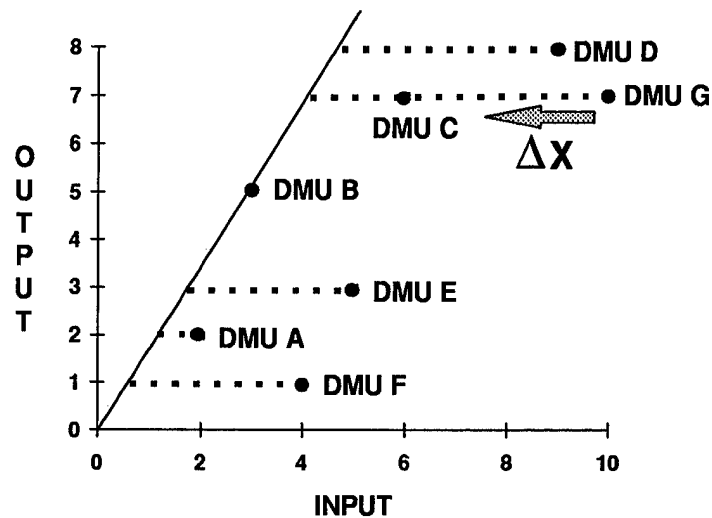


Figure 2.6: DEA Input Oriented CCR Model Envelopment Surface

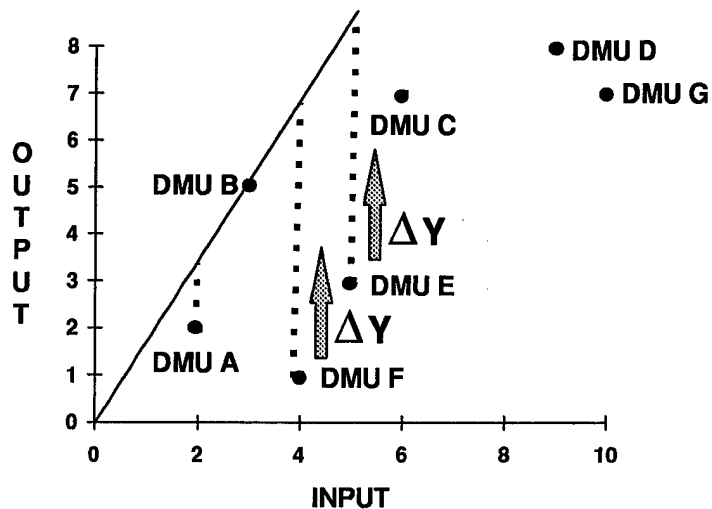


Figure 2.7: DEA Output Oriented CCR Model Envelopment Surface

The output oriented CCR model depicted in Figure 2.7 allows for an intuitive explanation of DMU efficiency. DMU F produces one unit of output using four units of

input and is inefficient. If DMU F were efficient, it would produce seven units of output given four units of input. Therefore, since DMU F produces only one seventh of what it should if it were efficient, DMU F has an efficiency score of 1/7 or .1429.

The mathematical formulation of the input oriented CCR model is:

$$\begin{aligned}
 &\text{Maximize: } z = \sum_r y_{rk} u_r \\
 &\quad \text{subject to} \\
 &\quad \sum_i x_{ik} v_i = 1 \\
 &\quad \sum_r y_{rj} u_r - \sum_i x_{ij} v_i \leq 0, \text{ for all } j=1, \dots, n \\
 &\quad u_r \geq \epsilon^* 1, \text{ for all } r \\
 &\quad v_i \geq \epsilon^* 1, \text{ for all } i \\
 &\quad \text{with} \\
 &\quad \epsilon \equiv \text{a non-Archimedean (infinitesimal) constant}
 \end{aligned}$$

The mathematical formulation of the output oriented CCR model is:

$$\begin{aligned}
 &\text{Minimize: } z = \sum_i x_{ik} v_i \\
 &\quad \text{subject to} \\
 &\quad \sum_r y_{rk} u_r = 1 \\
 &\quad - \sum_r y_{rj} u_r + \sum_i x_{ij} v_i \geq 0, \text{ for all } j=1, \dots, n \\
 &\quad u_r \geq \epsilon^* 1, \text{ for all } r \\
 &\quad v_i \geq \epsilon^* 1, \text{ for all } i \\
 &\quad \text{with} \\
 &\quad \epsilon \equiv \text{a non-Archimedean (infinitesimal) constant}
 \end{aligned}$$

The non-Archimedean (infinitesimal) constant is used as a lower bound for the virtual multipliers in the dual formulation (16:32).

The Banker, Charnes and Cooper (BCC) model (1984) results in a piecewise linear, VRS envelopment surface. Similar to the CCR model, the BCC model may also be input-oriented or output-oriented (16:43) as depicted in Figure 2.8 and Figure 2.9.

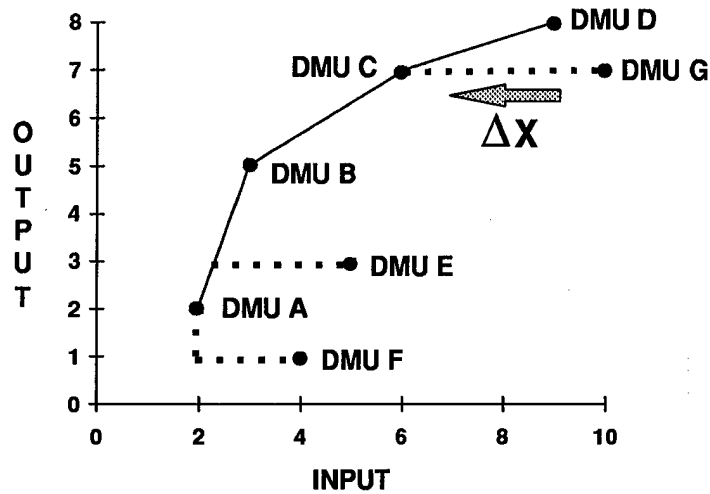


Figure 2.8: DEA Input Oriented BCC Model Envelopment Surface

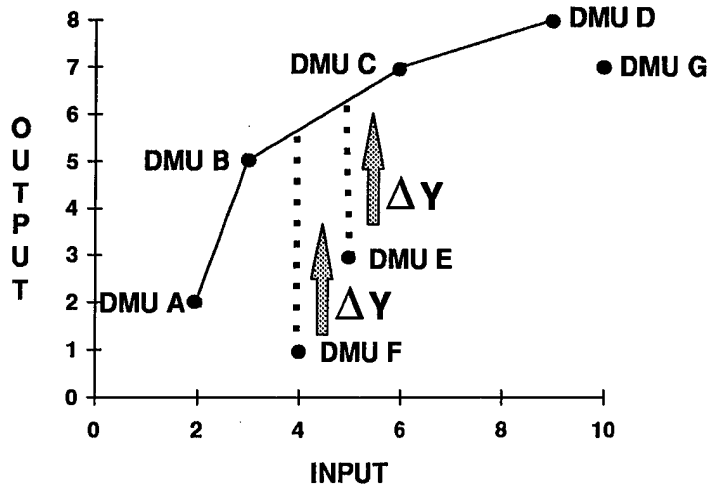


Figure 2.9: DEA Output Oriented BCC Model Envelopment Surface

As with the output oriented CCR model, the output oriented BCC model depicted in Figure 2.9 also allows for an intuitive explanation of DMU efficiency. DMU F produces one unit of output using four units of input and is inefficient. If DMU F were efficient, it would produce six units of output given four units of input. Therefore, since DMU F produces only one sixth of what it should if it were efficient, DMU F has an efficiency score of 1/6 or .1666.

The mathematical formulation of the input oriented BCC models is:

$$\begin{aligned}
 &\text{Maximize: } z = \sum_r y_{rk} u_r + u_o \\
 &\quad \text{subject to} \\
 &\quad \sum_i x_{ik} v_i = 1 \\
 &\quad \sum_r y_{rj} u_r - \sum_i x_{ij} v_i + u_o \leq 0, \text{ for all } j=1, \dots, n \\
 &\quad u_r \geq \epsilon^* 1, \text{ for all } r \\
 &\quad v_i \geq \epsilon^* 1, \text{ for all } i
 \end{aligned}$$

The mathematical formulation of the output oriented BCC model is:

$$\begin{aligned}
 &\text{Minimize: } z = \sum_i x_{ik} v_i + u_o \\
 &\quad \text{subject to} \\
 &\quad \sum_r y_{rk} u_r = 1 \\
 &\quad - \sum_r y_{rj} u_r + \sum_i x_{ij} v_i + u_o \geq 0, \text{ for all } j=1, \dots, n \\
 &\quad u_r \geq \epsilon^* 1, \text{ for all } r \\
 &\quad v_i \geq \epsilon^* 1, \text{ for all } i
 \end{aligned}$$

Each of the classic DEA models may also be programmed as an efficient or “super-efficient” formulation. The concept of super-efficiency was developed by Andersen and Petersen as a method to further discriminate among efficient DMUs (1:1262). By

eliminating the DMU under evaluation from the constraint set in the linear program, the evaluated DMU may attain an efficiency score greater than 1. The super-efficiency score represents the allowable percentage increase of resource use by that DMU which will still allow the DMU to remain efficient without a corresponding increase in output (Figure 2.10). For example, a DMU with an efficiency score of 1.25 can use up to 25% more resources to produce the same amount of output and it will still remain efficient compared to other DMUs.

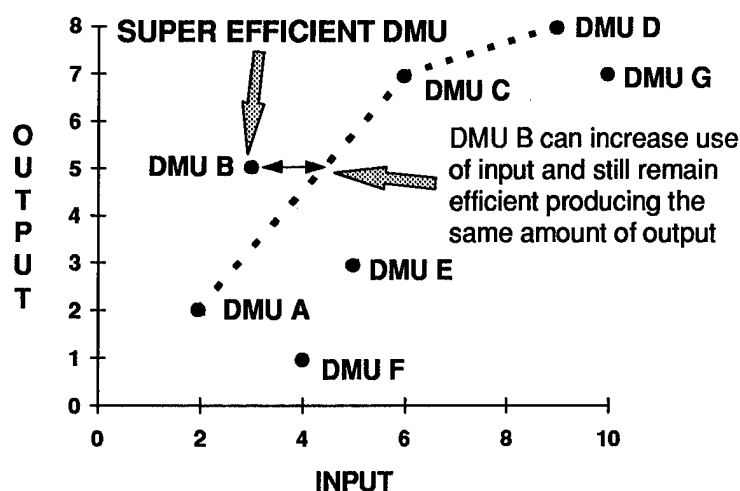


Figure 2.10: DEA Super-Efficiency Model Envelopment Surface

DEA models are a powerful tool because the analyst may compare vastly dissimilar but common resource inputs-- labor, capital, time, facilities, or environment-- without having to use the same quantifiable metric. Although multiple outputs are not discussed in this paper, DEA models can be used to determine the efficiency of firms which produce multiple outputs. As such, DEA is classified as a multiple criteria decision analysis method. DEA models are extremely useful for measuring the relative efficiency of organizations in the public/not-for-profit sector with multiple significant attributes or

where measurable parameters for evaluation are available, but are not usually expressed in dollar terms. DEA techniques have been successfully used in evaluating the managerial performance and efficiency of banks (9), the efficiency of schools (10), airlines (20), military units (50), and hospitals (6).

However, DEA models are not without their limitations. First, because DEA models are commonly considered deterministic models (30:311) and are not statistical in nature, there are no generally accepted statistical tests to determine the accuracy of the DMU's efficiency rating (15:54). An analyst can not be certain a particular DMU's efficiency rating of 0.99 is statistically different from another DMU's efficiency rating of 1.

Although recent work by Banker (4:1265) attempts to develop the statistical foundation for DEA, most of the literature to date is inconclusive. Banker proved that the DEA efficiency ratings are consistent, maximum likelihood estimators, and that their bias approaches zero for large sample sizes (16:111). Banker suggests specific hypothesis tests for the DEA estimators based upon the assumption that the error terms are distributed with an exponential or half-normal distribution. However, Banker's hypothesis tests rely on strict assumptions and he proves his hypothesis concerning the consistency of the DEA estimators only for the restrictive multiple input, single output scenario-- the focus of this particular research (42:441).

Additionally, because DEA models are usually assumed to be deterministic, we make the implicit assumptions that there is no random error in the data and the empirical efficient frontier is non-stochastic. If input or output data are actually stochastic random variables or estimates of stochastic variables, estimates of DMU efficiency or the

estimates of DMU virtual multipliers (what we are concerned with estimating for use in individual production functions in the FAARR model) may be subject to input data errors (42:442).

Since DEA models are non-parametric and data based (empirical), we must have a sufficient number of DMUs compared to the number of input and output variables in order to conduct the evaluation. If we have as many DMUs as input and output variables, there is the possibility all DMUs will be rated as efficient and the model will not be able to discriminate between efficient and inefficient DMUs. Charnes suggests using at least three times as many DMUs as inputs and outputs (21:621).

Finally, depending upon the homogeneity of the set of DMUs and the model's resource constraints, DMUs may choose feasible but highly unrealistic "weights" which maximize their efficiency. DMUs at the edge of the production possibility frontier or DMU's which utilize resource inputs significantly different from the average DMU are sensitive to this issue. These outlier DMUs' efficiency scores would rely heavily on relatively large "weights" for one or two specific inputs or outputs. Although these units may appear technically efficient and may lie on the production possibility frontier, their choice of virtual multipliers may make them allocatively inefficient (11:2), as illustrated by DMU D in Figure 2.3. Using prior knowledge or expert opinion concerning resource utilization and judiciously constraining the range of the DMU's virtual multipliers in the linear program, an analyst may derive a more realistic empirical efficiency frontier (16:54).

2.3 Deterministic versus Stochastic DEA Models

The majority of the operations research and management science community classifies the classical DEA techniques as deterministic models (47:109). DEA efficiency ratings are calculated assuming only a concave, monotonic functional form computed solely from input and output data. In the classical DEA models already discussed, there is no assumption as to the existence or distribution of an error term. Any deviation from the calculated efficient frontier is assumed to be due to inefficiency of the DMU and not due to stochastic noise or measurement error in either the input or output data. As such, this type of DEA model may be considered deterministic in the broadest sense of the term (42:442). Abraham Charnes, the co-inventor of the DEA methodology, states (21:621):

Every DEA analysis involves sample data of inputs and outputs which are converted by definite mathematical operations into other quantities. By definition such quantities are "statistics". Therefore *every* DEA model is a stochastic model. Since, however, the distribution functions of managerial performance at the different DMUs is unknown, we lack appropriate statistical theory for our real statistical structures.

The assumption of a deterministic DEA model generates the requirement to verify the accuracy of all input data. If the input data contains either random error or measurement error, the estimated production frontier or "efficiency surface" would be subject to stochastic perturbations and be biased upward (43:124). There is also the possibility that truly efficient DMUs--which actually determine the efficient frontier--are estimated as inefficient using DEA due to stochastic error in the data. Using the assumption of a deterministic DEA model, any estimates of the production frontier are vulnerable to outliers and measurement errors. If the analyst suspects stochastic input data, as a minimum he should conduct a thorough sensitivity analysis of the efficiency estimates

using the specific techniques discussed in the next section. The specific effect on the DEA analysis depends upon the stochastic nature of the system. If the system possesses minimal random error, the DEA derived empirical efficiency frontier should closely approximate the true frontier. Individual DMUs which are close to the empirical frontier may be considered efficient for all intent and purposes.

There has been some recent work attempting to blend the elegance and simplicity of the deterministic DEA model to the realities of the stochastic nature of data inputs. Charnes et al. (21) suggest window analysis where a number of DEA estimates are made for a set of DMUs over multiple time periods. They suggest that this technique not only indicates the stability of the DEA efficiency estimates, but may also reveal the nature of any stochastic variability. Using this technique, a more accurate estimate of DMU efficiency would be the median efficiency score for a particular DMU over all time periods or "windows" (21:622). This technique assumes the stochastic portion of efficiency errors are random with respect to time. Although this technique does not identify the source of the variability in the DEA efficiency estimates or hypothesize the probability distribution of the DEA efficiency scores, it is a useful tool to determine the stability of the DEA scores.

Sengupta suggests a number of data screening techniques to filter contaminated data for probable outliers based upon classical statistical tests. Since we do not know the true underlying distribution of input and output data, he suggests editing both input and output data using the non-parametric bounds of Chebyshev's inequality (44:17). Any DMU which has outlying data is not used in the calculation of the efficient frontier. Once we estimate the virtual multipliers for the remaining, "standard" set of DMUs, we can

calculate the efficiency of the outlier DMUs. Care should be taken any time we edit our data set or restrict the value of the virtual multipliers for any DMU--as is the case in the formulation of the current FAARR model.

Banker takes a more formal approach to the problem of stochastic variability and suggests "Stochastic DEA". Synthesizing the classical DEA model with goal programming, Banker decomposes a hypothesized error term into pure error and DMU inefficiency error. The pure or random error is considered symmetric with some unknown distribution. The DMU inefficiency error is positive, ensuring that inefficient DMUs fall below the efficient production frontier. As Stolp explains in his article, the Stochastic DEA model requires the analyst to assume a specific percentage of the total error is due to inefficiency and a specific percentage is due to random error. The analyst may also conduct sensitivity analysis of the DEA efficiency scores for different assumed percentages of the pure error term (47:110-111).

Olesen and Petersen confront the possibility of the stochastic nature of DEA efficiency scores by developing Chance Constrained Efficiency Evaluation (CCEE) (42). Similar to Banker's approach, CCEE assumes that the total error is composed of some percentage of pure error and the remainder of the total is due to DMU inefficiency. The CCEE technique is based upon chance constrained programming. Using a series of observations, the model estimates a confidence region for the efficiency estimate for each DMU. The CCEE model transforms the set of probability constraints into a set of deterministic constraints. The CCEE model requires a series of data for each set of DMUs. Additionally, there is an implicit assumption of no technical progress during the time

period of the data series. Any improvement to technical efficiency during the data time series would be erroneously decomposed into both true improvement and stochastic error.

Finally, Thomas suggests identifying a Robustly Efficient Comparison Set (RECS) of efficient DMUs (50). The dual variables in DEA contain a wealth of information concerning the production possibility frontier. Specifically, for each inefficient DMU, the dual variables identify efficient units most like the evaluated inefficient units. By identifying which efficient units are consistently identified by inefficient units, the analyst can determine a RECS--similar to identifying consistently best performing units across time using window analysis. Efficient units who are not used as common reference sets may only be efficient due to technical-- and not allocative-- efficiency. Similarly, these DMUs may have efficiency scores of 1 due to the stochastic error of some input or output variable. Use of the RECS may help to alleviate DMU misclassification-- characterizing an efficient DMU as inefficient or an inefficient DMU as efficient-- due to the stochastic nature of the variables (21:672).

2.4 DEA Sensitivity Analysis

Sensitivity analysis for mathematical programming techniques can be considered analogous to statistical testing for classical statistics techniques such as regression. Both methodologies are concerned with determining the range of allowable variation in the data. With linear programming, the analyst uses sensitivity analysis or parametric analysis to determine a range on the input variables or estimated coefficients where the optimal solution's basis does not change. In regression analysis, the analyst is concerned with

determining the range of values for estimated coefficients in which the hypothesized linear relationship remains statistically significant (17:139). As already mentioned, the literature has still not addressed the statistical theory for DEA and specific tests for the statistical significance of efficiency scores. Thus, we must turn our attention to sensitivity analysis in order to examine the stability of DEA estimates.

DEA requires the analyst to formulate and solve a linear program for each DMU. Because of the number of linear programs, the number of input and output variables, and the variation in the inverse matrix due to output changes for any one DMU (17:140), traditional sensitivity analysis for DEA models quickly becomes intractable. Because DEA is a descriptive tool, most analysts are not interested in the range of efficiency ratings estimates for a particular DMU. Most analysts are only interested in the relative change of a DMU's efficiency score versus other DMUs, or when a DMU no longer has an efficiency score of 1 and is no longer considered efficient. In recent years, researchers have developed many heuristics and techniques to discriminate between "robust", truly efficient DMUs and DMUs with unrealistic virtual multipliers.

Valdmanis suggests a simplistic, qualitative approach to determine the sensitivity of the DEA efficiency scores. He suggests initially conducting the DEA analysis and then systematically varying the number of input variables or selecting alternate input variables and recomputing the DMU efficiency estimates. In this manner, the analyst can observe the changes in the DEA efficiency estimates for each DMU assuming different resource input mixes and data sets. The truly robust and efficient DMUs should remain efficient for most resource combinations (52:195). In essence, there is probably only one or two

correct DEA model formulations of inputs and outputs which accurately estimate the efficiency of a set of DMUs.

Boussofiane et al., suggest the use of cross efficiency matrices. This technique indicates how a DMU's efficiency score is rated by other DMUs. The analyst constructs a matrix of DMU efficiency ratings using the virtual multipliers ("weights") of all other DMUs and then calculates the average efficiency score for each DMU. A DMU with a relatively high average efficiency score using the virtual multipliers from other DMUs is probably an efficient DMU. A truly inefficient DMU that appears efficient would have a high efficiency score using its own virtual multipliers. However, once the truly inefficient DMU uses the virtual multipliers of other DMUs, the truly inefficient DMU may no longer appear efficient (11:5).

Similarly, Charnes et al., (19) suggest imposing restrictions on the values of the virtual multipliers, or weights, which the linear program calculates for each input and output. Using prior knowledge of efficient operating practices or known physical limitations, the analyst constrains the values of the virtual multipliers in the linear program. An inefficient unit that was choosing an unrealistic or inappropriate range of values for its virtual multipliers would now appear less efficient. For example, in a manufacturing context, let us assume that we know through experimentation or historical data that a production process exhibits CRS and it takes three units of labor and one unit of capital to efficiently produce each unit of output. For one unit of output, we require three units of labor and one of capital. For two units of output, we require six units of labor and two of capital, and so forth. The historical, relative value or efficient ratio of capital to labor is three to

one. If a DEA model assigns a virtual multiplier to labor that is an order of magnitude different than historical efficient practice, we may judiciously constrain the value of the virtual multipliers in the DEA model to determine more realistic efficiency score estimates.

Thompson et al. in Charnes et al. (16) use Strong Complementary Slackness Conditions (SCSC) to conduct sensitivity analysis of DEA estimates for both farming and coal mining. By analyzing the dual variables of the linear program's optimal solution for each DMU, Thompson et al. determine the allowable data variation in the inputs and outputs which does not change the efficiency score of the DMU (16:397). They show that the efficiency ratings for these specific efficient DMUs are robust. To operationalize their sensitivity analysis methodology, Thompson et al. suggest varying a specific input or output vector for all DMUs by $\pm 5\%$ in a stepwise manner. The extreme efficient DMUs-- the subset of all efficient DMUs which are truly efficient-- would remain the most efficient throughout this stepwise process. The analyst would then have more confidence that the identified extreme efficient DMUs are the truly efficient DMUs. The specific DMU efficiency ratings become sensitive to the data variation when there is a change in the rank order of one DMU versus another.

Finally, Jaska formalized the mathematical theory underlying the approach of the SCSC and developed a more rigorous sensitivity analysis methodology called the Radius of Classification Preservation (RCP) for use with an additive DEA model. Using the L-1 and L-infinity norms as metrics, Jaska develops a linear programming formulation which estimates the minimum radius of a sphere or "ball" in n-space centered on the DMU's efficiency estimate. All input/output vectors contained within this ball are feasible

resource mixes where the DMU's estimated efficiency score would not change. The radius of this ball in n -space is the radius of stability. Using the radius of stability an analyst can compute the minimum change required in any input/output vector combination to change the estimated efficiency score of the specific DMU (32:94-95). This radius of stability may be calculated for both efficient and inefficient DMUs.

As just summarized, most of the literature and techniques for conducting DEA sensitivity analysis have been focused on the changes in the individual DMU efficiency scores, and not on the change in the value of the DEA virtual multipliers. Because DEA is a descriptive tool, the operations research community has been more concerned with the sensitivity of the ordinal ranking of the efficiency estimates versus the sensitivity of the cardinal values of the virtual multipliers. However, because this research evaluates the use of DEA efficiency estimates in a prescriptive, resource allocation model, we are concerned with the sensitivity of the cardinal values of the efficiency scores.

2.5 Beyond the Basic DEA Model

DEA models were originally developed solely for the purpose of efficiency evaluation (28:1173). In recent years, researchers have attempted to use the wealth of information provided from the basic DEA model in other mathematical programming and statistical models. For the most part, all of these multiple stage, mathematical models possess a common theme--use of information from a first stage, descriptive DEA model in the following stages of a prescriptive mathematical model. Commonly, these second stage models use some type of linear or L1 norm regression or a goal programming variant to

estimate parameters for an industry with a specific, hypothesized functional form. The focus of this research, the current FAARR model, is also a two-stage, descriptive/prescriptive DEA model.

Lovell, Waters, and Wood (16:329) used a modified DEA and regression based approach to construct a stratified model of education production in the short, medium, and long term. Their stratified model used the logarithm of secondary school super-efficiency scores as the dependent variable in a regression model. The second stage regression model provided statistically testable, estimated parameters which explained the variation in the schools' DEA scores. Using this information, the authors concluded that schools perform better meeting their medium and long term objectives and that there was greater room for policy decisions to impact the short term level of education production.

Bardhan, Cooper, and Kumbhakar also used a joint DEA/regression based model to estimate parameters for a production function first using DEA to identify efficient and inefficient units (8). This efficiency information was subsequently used in a regression model with indicator variables for the two populations of DMUs. Using a simulation model of a known production function, the authors concluded that classical statistics based techniques were not able to accurately estimate the true parameters of the production function. However, when the efficiency information from DEA was used in conjunction with regression based techniques, the estimated parameters for the efficient production function were statistically accurate.

Thomas also used a two stage DEA and goal programming model to estimate the parameters of an industry-wide, efficient production function for US Army Recruiting

battalions (50). Using a Multiplicative DEA model and facet analysis, the author identified a Robustly Efficient Comparison Set of DMUs. These efficient DMUs were then used in a goal programming model to estimate the parameters for a frontier production function. The estimated parameters from the model were used to conduct sensitivity analysis and estimate the marginal returns of varying levels of resources. This specific USAREC model was known as the FAARR-SHARE model. Charnes et al. used a similar model to estimate parameters for an efficient parametric production function for the Latin American airline industry (20).

Golany and Yu developed a goal programming-discriminant function to estimate an empirical production function based on DEA results (29). The authors identified efficient DMUs using an additive DEA model and then used a goal programming model to estimate the parameters of a Translog discriminant function. The discriminant function selects a separating hyper-plane which both segregates the inefficient and efficient DMUs into two groups and attempts to maximize the distance between the two respective groups. Golany and Yu then conducted a simulation analysis with a known production function in an attempt to demonstrate that their two stage model could outperform regression based techniques in retrieving the original parameters of the production function. Although their discriminant goal-programming model did out perform the regression based approaches, it was not able to accurately estimate the parameters for the known production function (29:181).

Two additional DEA models deal with the allocation of resources at the macro or industry level. These models may be classified as DEA-Resource Allocation Models

(DEA-RAM). First, Golany, Phillips, and Rousseau suggest reallocating resources at the macro level by constructing a mathematical program using DMU effectiveness indices to prioritize the allocation of resources (27). The effectiveness indices represent a DMU's efficiency transforming inputs into outputs compared to the average DMU. The effectiveness indices are computed for each DMU for each input and each output. The mathematical program's objective function uses the DMU's efficiency score and effectiveness indices to weight the allocation of resources between DMUs. The authors use an empirical example that demonstrates, in most cases, efficient DMUs were allocated increased resources which were proportionately taken from inefficient DMUs (27:8-9).

The second DEA-RAM model, developed by Golany and Tamir, uses a single mathematical program which combines an Additive DEA model with a weighted penalty function. The penalty function incorporates three competing objectives of efficiency, effectiveness, and equality in the allocation of resources (28). The authors define efficiency in the classical DEA context. Effectiveness is defined as the ability to produce some percentage of output given a fixed resource input--say graduate at least 85% of the school population. Equality is defined as the percentage change in a particular resource for a particular DMU from current levels to the levels prescribed by the DEA-RAM model. This objective ensures no DMU is allocated an inordinate amount of some resource at the expense of another DMU. Golany and Tamir demonstrate their DEA-RAM model using a simulation of a known production function.

As the DEA literature suggests, the DEA methodology is well developed, documented, and regarded within the Operations Research and Economics community as a descriptive

analytical tool. However, as a prescriptive tool, DEA based models are still in their infancy. The use of a non-parametric technique such as DEA to derive parameters for specific functional forms may be inappropriate due to model misspecification and the possible stochastic nature of input and output variables. We must remember that the parameters' estimates derived from DEA models represent a single observation from a single DMU for a specific time period and may not be indicative of the true, long term nature of the production frontier for all DMUs.

2.6 FAARR Model Background

As previously mentioned, the United States Army Recruiting Command (USAREC) contracted the Center for Cybernetic Studies at the University of Texas at Austin to develop the Forecast and Allocation of Army Recruiting Resources (FAARR) decision support system. The model was developed to provide USAREC with a rapid response methodology to forecast active Army high quality Graduate Senior Male Alpha (GSMA) contract production given fixed levels of resources, or forecast required resource levels given a fixed goal of GSMA enlistment contracts (13:5).

Traditionally, GSMA contracts are the hardest to recruit and require the most resources-- in both recruiter time and bonus or college incentives-- per contract compared to other lower quality recruits (13:5). Because the Army's Recruiting Command (USAREC) is organized into five brigades with 41 battalions, there are 41 DMUs in the FAARR model. Each DMU represents a separate battalion with a specific assigned

geographic area based upon the density of the DMU's client population and political boundaries.

The FAARR model uses a two stage DEA/optimization routine as depicted in Figure 2.11. The DEA virtual multipliers for each of the 41 recruiting battalions are calculated using linear programming within General Algebraic Modeling System software (14). The GAMS DEA model uses a Multiplicative, super efficiency, dual DEA model formulation. The virtual multipliers from the GAMS DEA model are used as estimated parameters for each battalion's production function in the FAARR model's second optimization phase. An EXCEL spreadsheet is used in the second phase optimization to forecast contract output or resource requirements given the DEA multipliers, resource levels, and market conditions.

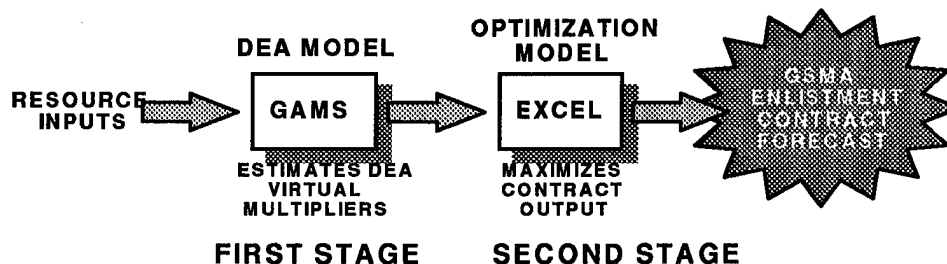


Figure 2.11: Army Forecast and Allocation of Recruiting Resources (FAARR) Model

The input and output data used in this research is quarterly data from 1st Quarter FY96 thru the 3rd Quarter FY97 and was supplied by the USAREC. The single DEA output is the number of GSMA contracts (GSMA). The eight DEA inputs are:

1. The number of On-station Producing Recruiters (OPR)
2. The national advertising Gross Rating Points (GRP) for broadcast (TVGRP)

3. The national advertising Gross Rating Points for radio (RADGRP)
4. The national advertising Gross Rating Points for print (MAGGRP)
5. The local advertising expenditures (LOCAL\$)
6. The number of Department of Defense (DoD) sister service recruiters (DODREC)
7. The unemployment rate (UNEMP)
8. The 17-21 year old male population (POP)

The general term “recruiting resources” is used to refer to all eight input variables.

In an economic or materials production paradigm, we can think of the population of a recruiting battalion’s area as the raw materials with which the recruiters (labor) will use their advertising dollars and GRPs (capital or, in a sense, factory machinery) to produce an output or product (GSMA contracts). The other two inputs-- competing DoD recruiters and local unemployment level-- define the competitive environment in which the Army recruiters work.

Several model variables are deterministic in nature. For example, the number of enlistment contracts is deterministic. The possibility of measurement or stochastic error is small. Similarly, the number of on station recruiters is also deterministic. Recruiters are intensely managed and monthly each battalion accurately reports the number of recruiters on formal unit strength reports.

However, four of the input variables are estimates of unknown actual values and are therefore stochastic--the 17-21 year old male population and Gross Rating Points for television, radio, and print. Area population data is supplied from commercial sources and is based upon forecasts using econometric models calibrated from the 1990 census.

Although the population forecast methods are very precise and generally accepted as accurate, they do possess a small amount of estimation variability.

Unfortunately, the Gross Rating Point (GRP) estimates for all media types are not as precise. These estimates are based upon sample Nielsen ratings obtained from families participating in the Nielsen research program throughout the United States. Nielsen ratings are obtained through the use of estimates and contain both sampling error and non-sampling error (41:43). Estimates of these distributions' variance for the historical data were not available from USAREC or the contracted advertising agency, Young and Rubicam. However, telephone conversations with Nielsen statisticians and the USAREC staff support an assumption that all estimates are normally distributed and accurate within plus or minus ten percent of the estimate.

The DEA model in the first stage of the FAARR model calculates the efficiency and derives the virtual multipliers for each of the 41 battalions. The model uses the natural logarithm of the input and output variable data and a linear program in the form:

$$\text{Maximize: } z = y_k w - \sum_i x_{ik} v_i + u_o \quad (1)$$

subject to:

$$y_j w - \sum_i x_{ij} v_i + u_o \leq 0, \quad j=1, \dots, n, \quad j \neq k \quad (2)$$

$$LB_i \leq v_i / \sum_h v_h \leq UB_i \quad i = 1, \dots, m \quad (3)$$

$$LB \leq w \leq UB \quad (4)$$

$$LB_{iw} \leq v_i / w \leq UB_{iw} \quad \text{for all } i \quad (5)$$

$$\sum v_i = 1 \quad (6)$$

where

$z \equiv$ estimated efficiency of battalion k
 $y_k \equiv$ natural logarithm of GSMA contract production (output) for battalion k
 $x_{ik} \equiv$ natural logarithm of recruiting resource (input) i for battalion k
 $w \equiv$ virtual multiplier for output
 $v_i \equiv$ virtual multiplier for input i
 $u_0 \equiv$ intercept term for battalion k
 $n \equiv$ total number of battalions being evaluated

This linear program is used n times to estimate the efficiency score for each battalion.

The objective function (1) calculates the recruiting battalion's technical efficiency by maximizing the difference between the weighted natural logarithm of the output and the weighted natural logarithm of the inputs. The literature refers to Equations (3) through (5) as linked cone constraints. Equations (3) and (4) simply constrain the range of the DEA virtual multipliers. Equation (5) constrains the value of pairs of virtual multipliers. Equation (6) normalizes the sum of the input variables' (recruiting resources) weights. This DEA model formulation is essentially a variant of the DEA VRS Multiplicative dual formulation with additional constraints on the virtual multipliers.

It is important to note the current GAMS DEA model mathematical formulation severely constrains the feasible values of the DEA virtual multipliers. Although individually constraints (Equations 3-6) may appear innocuous, combined they are very restrictive. This issue is analyzed in detail in Chapter 3.

Once the analyst uses the GAMS program to calculate the DEA virtual multipliers and efficiency scores for the battalions, this data is entered into the second stage EXCEL spreadsheet model. This second phase optimization model uses the DEA virtual multipliers to estimate a separate Cobb-Douglas production function for each of the 41 DMUs. The spreadsheet model has three modes. The analyst can:

1. Forecast enlistment contract production given a fixed number of recruiters and advertising dollars
2. Estimate the minimum required number of recruiters to meet a specified enlistment contract goal given a fixed number of advertising dollars.
3. Estimate the minimum amount of advertising needed to meet a specified enlistment contract goal given a fixed number of recruiters.

For this research, the author evaluated the model's first mode--forecasting the number of enlistment contracts given a fixed number of recruiters and advertising dollars.

The second phase EXCEL spreadsheet optimization model allows for allocation of the annual advertising budget to each quarter, battalion, and type of advertising medium based upon user preferences and the historical resource allocation for each recruiting battalion. The stated objective is to optimize the number of contracts (or mission) assigned to each recruiting battalion in order to maximize the total contract production across USAREC constrained by the individual recruiting battalion's calculated efficiency and virtual multipliers. It is important to recognize that if a battalion is only 80% efficient and we increase resources to that battalion, the model assumes the battalion will produce more output, but only at its estimated 80% efficiency. This recognizes the assumption that battalions with less than efficient performance will continue to perform in that manner (13:5). Additionally, the EXCEL model in this mode does not re-allocate resources from less efficient to more efficient battalions in order to maximize its forecasts. The model allocates total recruiter and advertising resources based upon a DMU's HISTORICAL PERCENTAGE of the total USAREC resources.

The optimization formulation for the EXCEL spreadsheet is:

$$\text{Maximize: } \sum_j y_j \quad (7)$$

subject to:

$$y_j w^* - \sum_i x_{ij} v_i^* + u_o^* = z^*, \quad j=1, \dots, n, \quad (8)$$

$$(1-B_j)Y_j \leq y_j \leq (1+B_j)Y_j, \quad j=1, \dots, n \quad (9)$$

where

$z^* \equiv$ DEA estimated efficiency of battalion k

$y_j \equiv$ natural logarithm of GSMA contract production (output) for battalion j

$x_{ij} \equiv$ natural logarithm of forecasted recruiting resource (input) i for battalion j

$w^* \equiv$ DEA estimated virtual multiplier for output

$v_i^* \equiv$ DEA estimated virtual multiplier for input i

$u_o^* \equiv$ DEA estimated intercept term for battalion j

$n \equiv$ total number of battalions being evaluated

$Y_j \equiv$ current contract production for recruiting battalion j

$B_j \equiv$ allowable percentage change from the current contract production for battalion j

The objective function maximizes the sum of the output for all 41 recruiting battalions. Equation (8) constrains the individual battalion to produce in accordance with its DEA determined production function from the first stage GAMS DEA model. Equation (9) constrains the forecasted output to remain within an arbitrary region. This constraint may be used to ensure a battalion does not receive a mission vastly greater than, or less than, its historic production. In its current formulation, the FAARR model does not contain equation (9). Including this equation may cause an infeasible solution for the mathematical program.

The specification of the FAARR model's second phase optimization program is similar to a more traditional, parametric, Cobb-Douglas efficient frontier benchmarking approach used by Horský and Nelson (31). Horský and Nelson derived their parameter estimates

(coefficients) not from DEA methods, but from a robust, Minimum Absolute Deviation (MAD) model which measures deviation from the efficient frontier. Horsky and Nelson estimated an efficient frontier sales production function for a sales firm with 230 salesmen organized in 26 separate sales districts. Their statistical and boot-strapped residual tests of the efficient frontier sales force production function provide anecdotal evidence for the correct specification of the FAARR model's second phase optimization routine.

III. Methodology

3.1 Introduction

As already discussed, the literature has not fully established the statistical foundation and specific probability distributions for DEA efficiency scores. Even if this theory were available, it would be of limited use in estimating the variance of total contract production using the FAARR model and estimating the subsequent confidence intervals for the model's GSMA forecasts. The FAARR model uses two separate, deterministic models to forecast total GSMA contract production. The DMU efficiency scores are just one of eleven statistics estimated in the first phase of the model. The DEA efficiency scores, in conjunction with the input and output virtual multipliers, determine the individual DMU production functions used in the second, optimization phase of the model. We not only need to explore the accuracy of the DEA efficiency scores, but we also require the distributions of the DMU virtual multipliers.

Most of the methodology and heuristics for conducting DEA sensitivity analysis assume the stability of the individual DEA efficiency scores. Existing theory does not address the sensitivity of the DEA virtual multipliers and how the virtual multipliers affect contract production forecasts.

If statistical theory for both the efficiency estimates and virtual multipliers were developed, calculating the variance of the forecasted contract production using statistics would still be intractable. The optimization routine used in the second phase of the FAARR model requires 451 estimates--eleven each for 41 DMUs. This includes the eight

input variables, the production function constant--or intercept--term, the individual DMU efficiency score, and the single output variable. In terms of a statistical model, this would require estimating, accounting for the interaction, and calculating statistical confidence intervals for 451 separate random variables. In short, at some future time an analytical or statistical solution may be possible, but may not be practical.

As stated, the first purpose of this research is to verify and validate the FAARR model and determine the accuracy and robustness of the model's forecasts. Because there are no analytical methods to estimate the standard error of the DEA derived parameter estimates, in order to validate the FAARR model, four primary tests of model accuracy and robustness were conducted.

1. The resource input and production output data set was analyzed.
2. Sensitivity analysis was conducted to determine the change in production forecasts due to changes in the DEA model virtual multipliers constraints and changes in aggregate recruiting resource levels.
3. Validation forecasts were made using three separate quarters of actual resource and production data to estimate the accuracy of the model's forecasts.
4. Finally, simulated data from a specified production function was used to determine the accuracy of the FAARR model's parameter and efficiency score estimates.

3.2 Input Data Analysis

The 1st Quarter FY97 data for the 41 DMUs was screened using Chebyshev's inequality in accordance with Sengupta's heuristic (44:17). This non-parametric test is

used to determine if the data set is homogenous and may be used to filter possibly contaminated data for outliers (44:17). Since an analyst may not have accurate information to identify the data set's probability distributions, Sengupta suggests using Chebyshev's inequality. It is evident from the informal screening of the data that the DMUs are not homogenous. The ranges on the resource inputs and contract outputs varied substantially, approximately ± 3 standard deviations from their means. Only four specific pieces of data-- OPRs, local advertising expenditures, TV GRPs, and the unemployment rate-- for three battalions were outside of the range of Chebyshev's inequality. These three recruiting battalions were not excluded from the data set because the information obtained from these data points could prove potentially useful. However, the high variance in the data indicated an increased probability an incorrect DEA model formulation would incorrectly classify DMUs as technically efficient. Since DMUs at the boundaries of the production set play a more important role in determining the empirical efficient frontier, any stochastic or measurement error may bias the efficient frontier estimates (44:17).

Kurskal-Wallis non-parametric tests were conducted on four recent quarters of data (3rd Quarter FY96 thru 2nd Quarter FY97). These tests indicated a rejection of the null hypothesis at the .05 level that the four different quarters of data came from the same distributions for the following variables: estimated DEA efficiency scores, GSMAs, all GRPs, local advertising dollars, and DoD recruiters. This indicates that there is an underlying trend in the data.

3.3 DEA Model Linear Programming Constraint and Resource Level Sensitivity Analysis

As already stated, most current research into the sensitivity analysis of DEA models is concerned with the sensitivity of the DEA efficiency scores and not the sensitivity of the DEA virtual multipliers. However, because both the virtual multipliers and the efficiency scores from the GAMS DEA model are used in the second stage, prescriptive optimization model, we must investigate these parameters' sensitivity to changes in the linear constraints of the DEA model. Similarly, we can investigate the change in the contract production forecast for changes in the aggregate amount of recruiting resources. For example, suppose we expect a "salami slice" 5% reduction in available recruiters, local advertising dollars, and national advertising due to budgetary constraints. What would be the corresponding percentage change in the FAARR model's GSMA contract forecast?

First, the sensitivity of the FAARR model forecasts to changes in the DEA model virtual multiplier constraints was analyzed. The assumption of an empirical Cobb-Douglas production function for the Army recruiting process in the second stage of the FAARR model translates into a distinct physical and economic interpretation. The parameter estimates for the Cobb-Douglas production function determine the output elasticities for that specific resource (18: 3747). The sum of the input resource elasticities determines if the industry is functioning at decreasing, constant, or increasing returns to scale (54:329). If the sum of the input resource elasticities is greater than one, the industry is IRS, if it is equal to one, the industry is CRS, and if it is less than one the industry is DRS. The ratio of the estimated input resource elasticities determine the Marginal Rates of Substitution (MRS) between resources (6:34). Thus, if the Cobb-Douglas estimated parameter for

population is 0.18 and the estimated parameter for the television GRP is 0.06, then the estimated MRS for one additional 17-21 year old is three television GRPs. Since the DEA virtual multipliers from the GAMS DEA model are explicitly used as Cobb-Douglas parameter estimates in the second stage EXCEL model, any DEA linear program constraint which limits the value of the virtual multipliers in the DEA model also limits the value of the Cobb-Douglas parameters in the EXCEL model.

Equations (3), (4), (5), and (6) are virtual multiplier constraints in the DEA linear program and limit the value of the eight input and one output virtual multipliers. Specifically, in the current FAARR DEA model Equation (3) constrains the sum of the virtual multipliers for all GRPs to be less than the virtual multiplier for Population (referred to from now on as Constraint Set 1) and constrains any input virtual multipliers to be less than three times any other input virtual multiplier (Constraint Set 2). Equation (5) also ensures the virtual multiplier for the contract output (GSMAs) is less than three times and greater than 1/3 of any other virtual multiplier (Constraint Set 3). Finally, Equation (6) constrains the sum of all virtual multipliers to equal one (Constraint Set 4).

As already stated, these constraint sets not only limit the value of the DEA virtual multipliers-- and limit the estimated efficiency scores for each DMU-- but they also limit the value of the estimated parameters used in the second stage Cobb-Douglas production function. For instance, by normalizing the input virtual multipliers and constraining their sum to be equal to one (Constraint Set 4), the model invokes a constant returns to scale for all inputs. However, the GAMS DEA model was formulated as a VRS Multiplicative model. Similarly, in the DMU parametric production functions used in the second stage

of the model, Constraint Set 2 limits the MRS for any two resources to be less than three. This means an additional recruiter (OPR) can be worth no more than \$3.00 in local advertising-- a dubious constraint at best.

Although restraining the virtual multipliers has a specific economic and managerial implication, the authors of the FAARR model provided no justifiable explanation for the arbitrary values of the virtual multipliers' constraint sets. A summary of the analysis of the total effect of the linked cone constraints is depicted in Table 3.1. The second and third columns indicate the feasible upper and lower bounds for each resource input variable. The fourth and fifth columns represent the actual bounds for the 1st QTR FY97 data set. As the table indicates, the range of values the input virtual multipliers may attain due to the sets of linked cone constraints is severely limited. Again, any virtual multiplier constraint has a specific economic interpretation and may result in invalid model estimates. The FAARR DEA model may be over constrained.

Table 3.1: Recruiting Resource DEA Virtual Multiplier Bounds

Resource	Theoretical	Theoretical	Actual	Actual
Input	Lower	Upper	Lower	Upper
Variable	Bound	Bound	Bound	Bound
Recruiters	0.0625	0.25	0.0714	0.25
Television GRPs	0.0556	0.1	0.0625	0.0838
Radio GRPs	0.0556	0.1	0.0625	0.0838
Print GRPs	0.0556	0.1	0.0625	0.0838
Local Advertising	0.0625	0.25	0.0625	0.25
DoD Recruiters	0.0625	0.25	0.0714	0.25
Unemployment Rate	0.0625	0.25	0.0714	0.2143
Population	0.1667	0.3	0.1875	0.2514

In order to estimate the sensitivity of the GSMA production forecast to changes in the virtual multiplier constraints and changes in recruiting resources, Constraint Sets 1, 2, and 3 were systematically relaxed one at a time. Then all virtual multiplier constraints, to include Constraint Set 4, were removed from the DEA linear program. The 1st Quarter FY97 data set was used as a baseline. GSMA contract production was forecasted with the second stage of the FAARR model and the calculated virtual multipliers using four different levels of recruiting resources: the actual recruiting resources for 1st Quarter FY97, a 5% increase in recruiters and television GRPs, a 10% increase in recruiters and television GRPs, and a 15% increase in recruiters and television GRPs. The increases in the two recruiting resources were not unrealistic scenarios given recent increases in total USAREC recruiter and advertising budgets.

The results of the sensitivity analysis are displayed in Table 3.2 and Table 3.3. Table 3.2 indicates the forecasted GSMA production and Table 3.3 indicates the forecasts' percentage change from actual production, referred to as the baseline. As Table 3.2 indicates, when the virtual multiplier Constraint Sets 1, 2, and 3 were relaxed, forecasted GSMA contract production changed. Without any virtual multiplier constraints or without the GSMA virtual multiplier constraint, the FAARR model could not find a feasible solution to the second stage linear program. Although the 1st stage DEA model can estimate efficiency scores for the DMUs, the 41 equality constraints in Equation (8) of the second stage, EXCEL spreadsheet model could not be satisfied unless forecasted contract production for certain DMUs was negative.

Table 3.2: GSMA Contract Forecast Sensitivity Analysis

FAARR DEA Model Virtual Multiplier Constraint Change	GSMA Forecast w/ change in OPR & TVGRP			
	None	+5%	+10%	+15%
Original FAARR model	7872	8497	9210	11739
GSMA Virt. Mult. <2* Any Input Virt. Mult.	7882	8599	11270	12297
GSMA Virt. Mult. <10* Any Input Virt. Mult.	7875	8605	13893	18005
No GSMA Virt. Mult. Constraint	Infeasible	Infeasible	Infeasible	Infeasible
Input Virt. Mult. < 4* Any Other Input Virt. Mu	7853	8492	10862	11770
Input Virt. Mult. < 8* Any Other Input Virt. Mu	7870	8591	11391	12499
No Input Variable Virt. Mult. Constraint	7917	8674	11621	12797
5*Population Virt. Mult. > Sum GRP Virt. Mult	7884	8613	11487	12658
No GRP Virt. Mult.Constraint	7889	8623	11522	12704
No Virtual Multiplier Linked Cone Constraints	Infeasible	Infeasible	Infeasible	Infeasible

Table 3.3: GSMA Contract Forecast Percentage Change

FAARR DEA Model Virtual Multiplier Constraint Change	% Change in Forecast w/ change in OPR & TVGRP			
	None	+5%	+10%	+15%
Original FAARR model	Baseline	7.94	17.00	49.12
GSMA Virt. Mult. <2* Any Input Virt. Mult.	0.13	9.24	43.17	56.21
GSMA Virt. Mult. <10* Any Input Virt. Mult.	0.04	9.31	76.49	128.72
No GSMA Virt. Mult. Constraint	Infeasible	Infeasible	Infeasible	Infeasible
Input Virt. Mult. < 4* Any Other Input Virt. Mu	-0.24	7.88	37.98	49.52
Input Virt. Mult. < 8* Any Other Input Virt. Mu	-0.03	9.13	44.70	58.78
No Input Variable Virt. Mult. Constraint	0.57	10.19	47.62	62.56
5*Population Virt. Mult. > Sum GRP Virt. Mult	0.15	9.41	45.92	60.80
No GRP Virt. Mult.Constraint	0.22	9.54	46.37	61.38
No Virtual Multiplier Linked Cone Constraints	Infeasible	Infeasible	Infeasible	Infeasible

As this sensitivity analysis indicates, the FAARR model forecasts are only slightly sensitive to the value of the virtual multiplier constraints in the DEA linear program as long as the forecasts use approximately the same level of recruiting resources as those used to estimate the DEA virtual multipliers. The maximum percentage change in total forecasted contract production was less than 1%. This analysis indicates the model is robust to changes in the virtual multiplier constraints assuming the forecasts use approximately the same recruiting resource levels as those used to compute the DEA virtual multipliers.

The FAARR model forecasts are more sensitive to changes in the virtual multiplier constraints when the recruiting resource levels deviate from the levels used to estimate the DEA virtual multipliers. Assuming a 5% increase in television GRPs and a 5% increase in

On-station Production Recruiters from 1st Quarter FY97 levels, using the currently formulated FAARR model, forecasted production increased by 7.94% compared to the baseline for a forecast of 8497 GSMA contracts². Using a 5% increase in recruiting resources in the 2nd stage of the model and relaxing the virtual multiplier constraint sets in the 1st stage DEA model, forecasted contract production increased by as much as 10% from the baseline forecast. The sensitivity of the forecasted production to changes in the virtual multiplier constraints quickly increased as recruiting resource levels changed from those used to estimate the DEA virtual multipliers. With a 15% increase in recruiters and television GRPs, relaxing the virtual multiplier constraints resulted in drastically increased forecasts-- from 49% to as much as 128% of actual production.

Additionally, the FAARR forecasts are also sensitive to changes in the aggregate level of all recruiting resources for every battalion. Again, using 1st Quarter FY97 recruiting resources and virtual multipliers as the baseline, the author varied all recruiting resources for all battalions by a fixed percentage from 75% to 125% of baseline levels. If the FAARR model could accurately forecast a CRS process, we would expect forecasted contract production to change in the same proportion as the aggregate resource levels. A 10% increase in resources for a CRS process would result in a 10% increase in forecasted contract production. This was not the case. As illustrated in Figure 3.1, a 5% decrease in all resources resulted in a 29.4% decrease in forecasted production. Similarly, a 5% increase in all resources resulted in a 42.74% increase in forecasted production. The forecasts were VRS, increasing throughout this range of resource variation. This simple

² The FAARR DEA model is formulated as VRS. A 5% increase in only two of eight recruiting resources

sensitivity analysis indicates FAARR model forecasts are extremely sensitive to changes in the resource levels from those used to calculate the DEA virtual multipliers and efficiency scores. The FAARR forecasts are not robust to varying levels of recruiting resources.

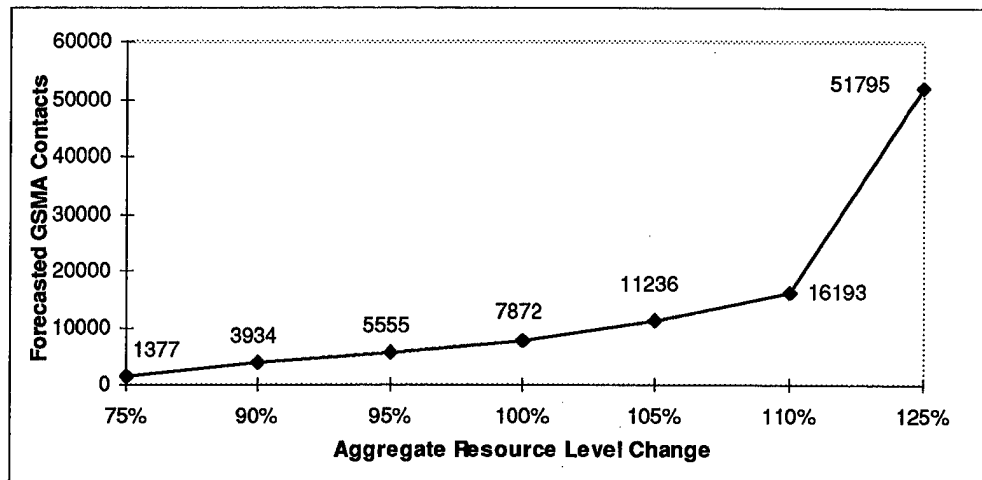


Figure 3.1: FAARR Model Resource Level Affect on Contract Production Forecast

It is also important to recognize that the current FAARR 2nd stage, EXCEL mathematical model formulation is similarly very restrictive and does not optimize. The feasible solution space for the mathematical program without equation (9) is exactly one point. The 41 efficiency constraints (8) for the linear program are equality constraints. Since the DEA virtual multipliers (w^* and v^*) and DMU efficiency scores (z^*) are determined from the GAMS DEA model, and the recruiting resource allocation (x_{ijs}) are based on a historical percentage of resources and user input, this linear program simultaneously solves the production function/efficiency constraints for the 41 battalions. This linear program can not actually optimize since there is only one unique solution to the program given any fixed allocation of resources. If the recruiting resources used in the

results in an 8% increase in forecasted production. A 10% increase in these two resources resulted in a 17% increase in forecasted production.

forecast are increased from their current levels and the battalions remain at their estimated efficiency, then contract production has to increase to satisfy the production function efficiency constraint (Equation 8). Similarly, if the recruiting resources used in the forecasts are decreased for all battalions, then contract production has to decrease to satisfy the production function efficiency constraint. If the maximizing function--

Maximize: $\sum_j y_j$ -- is replaced with the minimizing function-- Minimize $\sum_j y_j$ -- the

mathematical program's solution would be the same. Again, FAARR's 2nd stage EXCEL spreadsheet optimization model can not optimize. It merely simultaneously solves the production function efficiency constraints (Equation 8) for the 41 recruiting battalions.

3.4 FAARR Model Validation Forecasts of Actual Contract Production

One qualitative method to measure forecast model accuracy and validation forecasting. This methodology can be described by the simple question: Can the model accurately predict the past? Validation forecasts, referred to as *ex post* forecasting (35: 209), allow the analyst to objectively measure the accuracy of a forecasting model by using the model to predict what has actually already occurred. If a forecasting model can not accurately predict the past, then it will probably not be able to accurately predict the future.

Forecasts for the first, second, and third quarter of FY97 were calculated using the actual resource allocation for the first three quarters of FY97 and the FAARR model virtual multiplier and efficiency score estimates from preceding time periods in FY96. Three estimates were made for each quarter using three different sets of DEA virtual

multipliers and efficiency scores. First, the DEA virtual multiplier and efficiency score estimates for the same quarter from the previous year were used. The authors of the FAARR model developed the model to use these estimates. Second, the DEA estimates from the immediately preceding quarter were used. Third, the average quarterly DEA estimates from the entire previous year were used. Using the DEA estimates from the immediately preceding quarter may induce undue seasonality into the FAARR forecast and result in a biased estimate for the following quarter. Non-parametric tests indicate some of the recruiting resource data may be seasonal. Similarly, using the average quarterly DEA estimates for the previous year may “smooth out” the seasonal component of the FAARR estimate and may again result in a biased estimate for that particular quarter.

As Measures Of Effectiveness (MOEs) to evaluate the various forecasting models, the overall model Mean Absolute Percentage Error (MAPE) for the sum of the forecasts for all 41 recruiting battalions, and the average and maximum MAPE across all 41 recruiting battalions was chosen. These statistics not only provide an indicator of overall model accuracy, but also express some measure of the variability of the estimates for each battalion. Table 3.4, Table 3.5, and Table 3.6 summarize the results of the validation forecasts.

Table 3.4: FAARR Model Forecast MOEs using DEA Virtual Multipliers from Same Quarter in Previous Year

Quarter	Overall MAPE	Average BN MAPE	Maximum BN MAPE
QTR 1 FY97	14	46	153
QTR 2 FY97	49	88	1,025
QTR 3 FY97	Infeasible	Infeasible	Infeasible

Table 3.5: FAARR Model Forecast MOEs using DEA Virtual Multipliers from Previous Quarter

Quarter	Overall MAPE	Average BN MAPE	Maximum BN MAPE
QTR 1 FY97	324,289	257,375	10,550,123
QTR 2 FY97	235	281	2,146
QTR 3 FY97	40	56	157

Table 3.6: FAARR Model Forecast MOEs using Average DEA Virtual Multipliers for all Four Quarters from Previous Year

Quarter	Overall MAPE	Average BN MAPE	Maximum BN MAPE
QTR 1 FY97	24	44	188
QTR 2 FY97	95	111	549
QTR 3 FY97	21	43	143

Analysis of the validation forecasts indicate none of the forecasts were adequate. The total overall model Mean Absolute Percentage Error (MAPE) from estimated versus actual production for the three forecasts ranged from 14.% to almost 3,500 times actual production. The best overall model MAPE was 14% for the 1st Quarter FY97 forecast, but this forecast's average battalion MAPE was 46% with a maximum MAPE of 153%. Although this specific model's forecast overall MAPE was a relatively small 14%, the individual MAPEs for each battalion varied greatly from actual contract production.

For comparison, Table 3.7 contains the forecasts for the Naive Forecast 1 model for the first three quarters of FY97. The Naive Forecast 1 simply forecasts the upcoming quarters production using the actual production from the previous quarter. The model does not account for trend or seasonality. In essence, the Naive Forecast 1 is not a forecasting technique at all, but merely uses as a forecast the most recent information available concerning the battalions' actual contract production (37:47). The difference in MAPE obtained from the Naive Forecast 1 and a more complicated forecasting model provides a measure of the improvement attainable using a more formal forecasting method (37:48). Although it is a simple forecasting technique, forecasts from the Naive Forecast 1 model were more accurate than the forecasts from the FAARR model for all MOEs.

Table 3.7: Recruiting Battalion Naive Forecast 1 Results

Quarter	Overall MAPE	Average BN MAPE	Maximum BN MAPE
QTR 1 FY97	3	16	47
QTR 2 FY97	8	12	46
QTR 3 FY97	17	19	71

3.5 FAARR Model Estimation of Production Function Parameters and DEA Efficiency Scores from the Simulation of a Known Production Function

Finally, the last technique used to evaluate the FAARR model involved the use of a simulation model to randomly assign efficiency scores and resource inputs to sets of simulated DMUs. These random inputs and efficiency scores were used with a known production function to calculate theoretical, known production output. The author then used the FAARR model to evaluate the set of randomly generated DMUs' inputs and outputs in an attempt to retrieve the actual production function parameters and DMU efficiency scores. Although DEA methodology is a descriptive tool and was not developed to explicitly estimate the actual parameters of a function, the FAARR model explicitly uses DEA estimated virtual multipliers as model parameters to forecast production. Thus, the author evaluated the FAARR model's ability to estimate function parameters and efficiency scores using a simulation of an actual production function. In

their article on estimation of empirical production functions, Golany and Yu state (29:174):

“...one way for a frontier estimation technique to prove its credibility is to demonstrate its ability to retrieve the elements of the original function. As a minimal requirement, it should be able to retrieve the correct ordinal ranking of the input elasticities. A more demanding requirement would be to test the estimated function with the original inputs and measure the distance between the estimated and theoretical outputs.”

The theory and methodology for using a simulation to measure the accuracy of a DEA model is depicted in an article by Banker, Chang, and Cooper (5). In their article, Banker, Chang, and Cooper compare the ability of Corrected Ordinary Least Squares (COLS), a translog function, and BCC and CCR formulated DEA models to estimate the true efficiencies of sets of simulated DMUs. The randomly selected amount of resource inputs for each DMU determined a theoretical “efficient” production using the Cobb-Douglas production function. This known efficient production was then multiplied by a random variable which decremented total production for 70% of the DMUs.

Applying a similar methodology to the FAARR DEA model, two simulation models of known production functions were constructed using GAMS software. Two different, Constant Returns to Scale (CRS), Cobb-Douglas production functions were used with random input resources selected from a multi-variate normal distribution estimated from actual 1st Quarter FY97 recruiting data. Model 1’s Cobb-Douglas production function had true parameters (coefficients) which were selected to not conform to the virtual multiplier linked cone constraints of the FAARR model. Model 2’s Cobb-Douglas production function had known function parameters which satisfied the fairly restrictive constraints of the FAARR model DEA formulation. The distribution of the random, actual or “true”, efficiency scores for DMU j is represented by the technical inefficiency term η_j , (46:236), where $\eta_j \in [0,1]$ and was selected from a truncated normal distribution

estimated from the FAARR model DEA scores using historic data³. The estimated normal distribution parameters were such that approximately 11.2% of the DMUs are efficient. No random error term was added to the model. Each simulation replication randomly selected a resource input and efficiency vector for 41 DMUs. One hundred simulation replications were conducted for each model. The mathematical formulation for the known production functions were:

$$y_j = \alpha_o \prod x_{ij}^{\beta_i} \eta_j \quad \sum_i \beta_i = 1$$

where

y_j \equiv output of simulated battalion j

α_o \equiv production function intercept term

x_{ij} \equiv input i for simulated battalion j

β_i \equiv coefficient for input i

η_j \equiv known efficiency for simulated battalion j selected from truncated normal distribution

These two models were deliberately constructed using specific random input variable distributions, known parameters, and no random error, to mirror and also favor the current FAARR DEA model formulation. The author hypothesized that by giving the FAARR model the “benefit of the doubt” with regard to the simulation model formulation, the FAARR model would accurately estimate the known simulated production function parameters. However, this was not the case.

³ The FAARR DEA model was not formulated as a super-efficiency model in these simulations. Theoretically, super-efficiency indicates the total percentage increase in resources for a particular DMU for which the DMU would remain efficient if it produced no more output. Since no battalion can have an actual efficiency greater than 1, a super-efficiency DEA model formulation would upwardly bias the Mean Absolute Deviation (MAD) efficiency estimate from its true value. Evaluating the accuracy of the FAARR model without the super-efficiency formulation will not change the efficiency scores for any inefficient DMU.

The actual and estimated average efficiency scores and parameters for each model are depicted in Table 3.8. Additionally, the correlation coefficient between the actual and estimated efficiency and the model's Average Percentage Error Rate (APER) for classification is listed. The APER in this context is similar to descriptive statistics used in discriminant analysis (34:230). The APER measures the relative number of times the FAARR model incorrectly classified a battalion as efficient when it was not efficient, or classified a battalion as inefficient when it was efficient. As these results indicate, the FAARR model was not able to accurately estimate the known, simulated production functions' parameters or efficiency scores.

Table 3.8: Simulation Model Results Using FAARR Model to Estimate Efficiency Scores and Parameters from a Known Production Function

VARIABLE PARAMETER	MODEL 1 KNOWN	MODEL 1 ESTIMATE	MODEL 2 KNOWN	MODEL 2 ESTIMATE
Intercept	1.2	5.115	0.18	5.2944
OPR	0.3	0.1596	0.18	0.1549
TV GRP	0.11	0.0665	0.06	0.0678
Print GRP	0.02	0.0665	0.06	0.0678
Radio GRP	0.06	0.0665	0.06	0.0678
Population	0.04	0.1994	0.18	0.2033
Local Adv \$	0.03	0.0878	0.18	0.0945
Unemp Rate	0.24	0.1577	0.18	0.1444
DoD Recruiters	0.2	0.1963	0.1	0.1997
Mean Efficiency Score	0.881	0.543	0.881	0.6421
Correlation Coefficient		0.13		0.16
APER		13%		14%

Not only were the average DEA estimated battalion efficiency scores drastically different from the actual simulated battalion efficiencies, the average correlation coefficient between each battalion's actual and estimated efficiencies for the Model 1 and Model 2 were only 0.13 and 0.16, respectively. Since DEA models measure the relative

efficiency of a set of battalions, we can not expect the estimated efficiencies to exactly correlate with the actual efficiencies. However, the FAARR model's efficiency estimates' correlation with the true efficiencies was strikingly low. In general, we can conclude a DEA model formulation which has a higher correlation coefficient than another DEA model formulation is a more accurate model (46:243). Additionally, the FAARR model was not able to accurately estimate the simulated production functions' known parameters. FAARR model parameter estimates highlighted in gray were significantly different than the parameters' true values.

A final Measure Of Effectiveness (MOE) is APER: the ability of each model to accurately identify the efficient and inefficient battalions. This is a common use of many DEA models. A simulated battalion is incorrectly classified if the DEA model classifies the battalion as efficient when it is not efficient or if the model classifies an inefficient battalion as efficient. The FAARR model incorrectly classified 13% and 14% of the simulated battalions for each production function, respectively. Table 3.9 and Table 3.10 present the confusion matrices for Model 1 and Model 2, respectively. The tables indicate the average number of the 41 simulated battalions incorrectly and correctly classified from all 100 simulation replications. For example, on average the simulation generated 4.18 efficient battalions and Model 1 classified 3.97 of these efficient battalions as inefficient. Similarly, on average the simulation generated 36.82 inefficient battalions and Model 1 incorrectly classified 1.65 of these battalions as efficient. Each model incorrectly classified efficient battalions as inefficient 95% of the time and incorrectly classified inefficient battalions as efficient 5% of the time. We would expect a VRS DEA model such as the

FAARR model to overestimate the efficiency of battalions generated from a CRS process. However, this was not the case. Average DEA estimated efficiency was less than the average true battalion efficiency. Again, it appears the DEA linked cone constraints may be overly constraining the values of the virtual multipliers, resulting in decreased DEA estimated efficiency scores and the majority of efficient battalions being classified as inefficient.

Table 3.9: Simulation Model 1 Confusion Matrix

FAARR DEA Classification	Actual Classification		Totals
	Efficient	Not Efficient	
Efficient	0.21	1.65	1.86
Not Efficient	3.97	35.17	39.14
Totals	4.18	36.82	41

Table 3.10: Simulation Model 2 Confusion Matrix

FAARR DEA Classification	Actual Classification		Totals
	Efficient	Not Efficient	
Efficient	0.2	1.81	2.01
Not Efficient	3.98	35.01	38.99
Totals	4.18	36.82	41

Finally, using the second simulated production function, the author increased the aggregate resource level for all battalions and used the FAARR model to forecast contract production. Since the simulation was a CRS production function, a 5% increase in resources resulted in an actual 5% increase in production. However, as depicted in Figure 3.2, the FAARR model forecasts were VRS and increasing throughout the range of resources used for the forecasts. A 5% increase in recruiting resources resulted in a 27.2% increase in forecasted production. This analysis indicates that if the actual production process is not VRS, FAARR model forecasts will not be accurate. The FAARR model incorrectly attributed the simulated battalions' less than efficient

production to a change in returns-to-scale and not to actual battalion inefficiency. This may be the cause of some of the inaccuracy in the FAARR model's forecasts when recruiting resource levels are increased.

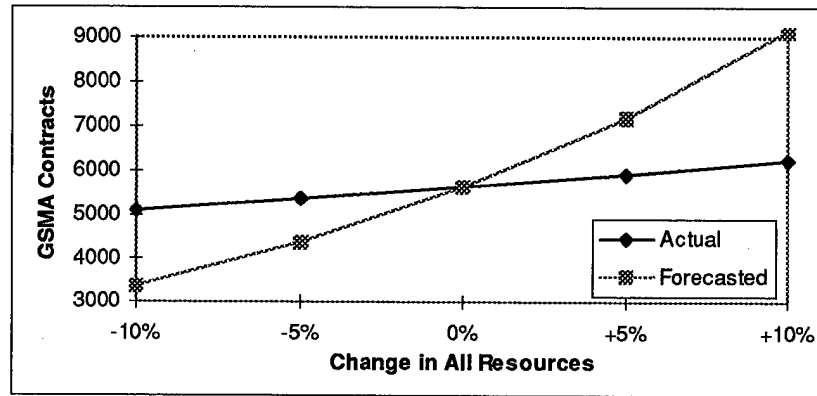


Figure 3.2 FAARR Forecast Evaluation of Simulated Function

3.6 FAARR Model Validation Summary

In summary, the author used sensitivity analysis, back forecasting, and simulation of a known production function to quantitatively and qualitatively estimate the accuracy and robustness of FAARR model forecasts. The sensitivity analysis demonstrated the FAARR model contract production forecast is sensitive to both the linked cone constraints of the GAMS DEA model and any changes in recruiting resource levels from those used to estimate the DEA model. Without any constraints on the virtual multipliers, the FAARR model could not find a feasible solution for the production forecast. Using actual 1st Quarter FY97 recruiting resources as a baseline, a relatively small 5% increase in the aggregate level of all recruiting resources resulted in a 42.74% increase in forecasted production. This analysis indicates the FAARR model may not be useful for “what if” analysis when forecasts’ recruiting resources change from their current levels.

The FAARR model was also not able to accurately forecast actual contract production for the first three quarters of FY97. Large forecast estimates' MAPEs were mainly due to large forecast errors for specific battalions, indicating the FAARR model can not accurately predict contract production for individual recruiting battalions.

Finally, using two simulation models of known production functions, the FAARR model was not able to accurately estimate the actual battalion efficiency scores, the actual production function parameters, or accurately classify battalions as efficient or inefficient. The FAARR model incorrectly classified efficient battalions as inefficient 95% of the time. Using data from one of the simulated CRS production function, a relatively small 5% increase in the aggregate level of all recruiting resources resulted in a 27.2% increase in forecasted production. Although estimating the underlying production function parameters is not necessarily important for a descriptive DEA model, the FAARR model was still not able to accurately discriminate between the simulated efficient and inefficient recruiting battalions. Additionally, the FAARR model assumes a VRS production process and efficiency scores and forecasts are calculated accordingly. If the actual underlying production process is not VRS, FAARR forecasts will not be accurate.

As already stated, DEA models are descriptive, non-parametric models-- they only indicate whether a DMU is efficient or inefficient. The DEA virtual multipliers are merely indicators of a resource's relative value which the linear program uses to ultimately arrive at an efficiency score. These virtual multipliers are the result of a single observation of DMU performance and may be influenced by stochastic error, measurement error, or seasonality. The FAARR model's second phase optimization routine explicitly uses these

virtual multipliers as parameters in a prescriptive mathematical forecasting model. As this analysis of the FAARR model indicates, use of numerical values from a non-parametric, descriptive model in a parametric, prescriptive model may be a dubious technique. Although the FAARR model may fit the existing recruiting data fairly well, attempts to forecast production using actual, but significantly different, levels of recruiting resources produced unacceptable results.

It seems in an attempt to use an unconventional mathematical model for both parameter estimation and forecasting, the FAARR model does neither very well. It may be more appropriate and more accurate for USAREC to use an Ordinary Least Squares (OLS) based econometric model to estimate specific resource parameters and a separate, time-series, Box-Jenkins or smoothing model for forecasting contract production.

Further, the results of this analysis indicate the FAARR first phase DEA model may be misspecified. The DEA model may have irrelevant variables, may not include relevant variables, may be formulated with an inappropriate envelopment frontier, or the linked cone constraints for the virtual multipliers may be too restrictive. The next section describes a strategy which may be used to select an appropriate DEA model and reduce the probability of model misspecification.

3.7 A Method for Selecting an Accurate DEA Model Formulation

Although the current FAARR model has some specific limitations, the wealth of information available from DEA models should not be discounted or discarded. This

section describes a three phase strategy for selecting an accurate DEA model formulation using Principal Component Analysis (PCA), Ordinary Least Squares (OLS) regression, and Monte-Carlo simulation. Efficiency information obtained by correctly identifying efficient and inefficient DMUs using an accurate DEA model formulation can then be used in other mathematical models to improve the accuracy of parameter estimates or contract production forecasts.

The advantages of this type of combined DEA/OLS model for production function estimation is outlined in an article by Bardhan, Cooper, and Kumbhakar (8). In simulation studies, use of DEA information to improve OLS models resulted in more accurate parameter estimates (8:25).

In order to identify the most accurate-- and thus the most appropriate-- DEA model formulation, the analyst not only needs to identify the appropriate input and output variables, but he/she also needs to identify the most appropriate form of the envelopment frontier--either Additive, BCC, CCR, Multiplicative, etc. Selection of a specific envelopment frontier also makes explicit assumptions concerning industry-wide returns to scale. For example, if a CCR formulation is used, a Constant Returns-to-Scale (CRS) process is assumed. Similarly, a BCC formulation assumes Varying Returns-to-Scale (VRS). Incorrect choice of the shape of the DEA envelopment frontier may lead to inaccurate efficiency estimates.

The three phase Principal Component Analysis / OLS / Monte-Carlo Simulation strategy is summarized as follows:

- First, use Principal Component Analysis and/or an OLS model to identify the relevant input variables.
- Second, use an OLS model or frontier estimation model containing the relevant input variables as an estimate for an appropriate “rough cut” production function.
- Third, use a simulation of this estimated production function to select the most accurate envelopment frontier for the DEA model formulation.

This specific DEA model formulation is used to identify efficient DMUs. The information provided from the accurate DEA model formulation can then be used with dummy variables in another OLS model to improve model fit and increase the accuracy of parameter estimates or forecasts (8:2). The strategy may be repeated if the new OLS model suggests including additional variables in the DEA model.

In the first phase of the strategy, Principal Component Analysis (PCA) and OLS are used as screening tools to “weed out” grossly inappropriate input variables for the subsequent DEA model. Although DEA models are not as prone to the deleterious effects of mis-specification as traditional statistical methods (47:112), care should be taken to not recklessly use all available data. Relevant input variables should be chosen on the basis of data accuracy, minimal data intercorrelation, and from data which is known to be related either “statistically, experientially, or conceptually” to the production process (16:427).

Given a fixed number of DMUs to evaluate, as the number of input and output variables increases, DEA can fail to discriminate between efficient and inefficient DMUs due to the increased dimensionality of the solution space (46:238). Inclusion of an irrelevant variable in a DEA formulation may also affect the model’s ability to discriminate

between efficient and inefficient DMUs. Including an irrelevant variable increases the number of constraints in the DEA linear programming formulation and therefore can not reduce, but may increase, estimated efficiency (46:241). A misspecified DEA model may rate a DMU which utilizes a small amount of an irrelevant resource as efficient. However, if the same DMU were to be evaluated in the correctly specified, lower dimensional space representing only the relevant input variables, that DMU may no longer be efficient. Conversely, omitting a relevant variable from a DEA model formulation may result in a reduction in estimated efficiency (46:239). In summary, too many input variables or an inappropriate choice of input variables may hide the true, efficient frontier.

Principal Component Analysis (PCA) and OLS can be used to screen the available input variables to select the most relevant subset of variables to include in the DEA model formulation. PCA is a factor analysis method used as a data reduction technique or as a tool to assess the underlying dimensionality of multi-variate data. PCA is commonly used to screen a set of potential variables for further statistical analysis (34:238). Analysts can use PCA to identify a less correlated (orthogonal) subset from a larger set of highly correlated data (34:238). Although the total number of variables may be reduced, much of the original data's variance information is maintained (23:23). The factor loadings matrix and eigenvalues for each factor are determined using fundamental matrix algebra with the input variable correlation matrix. Any input variable which does not highly load on the most important factors-- as determined by the largest eigenvalues-- may not be a relevant or significant variable to subsequently use in a DEA model.

Similarly, the analyst can conduct initial variable screening using step-wise OLS procedures. The analyst could include variables in the DEA model which may be significant in the OLS model at only the .15 or .25 level, but are not statistically significant at the .05 or .10 level. The variables included in the DEA model may not be statistically significant at an appropriate level in the OLS model, but they may be pertinent nonetheless--as proven by past experience, theory, or expert opinion.

Using the relevant variables selected from the PCA and OLS screening, the analyst constructs an approximate OLS or efficient frontier model of the production function. This mathematical model is an approximation of the true relationships between the input and output variables.

There is some discussion in the literature on the semantic definition of a production function. If we simply define a production function as the approximate mathematical relationship between a set of input variables and a set of output variables, then OLS regression techniques can be used to estimate this input to output relationship. Deviations between the actual and estimated function may be both positive and negative and these deviations can be attributed to inefficiency, stochastic error, or units producing at more than 100% efficiency. However, most authors define a production function in a more exact and restrictive manner. Production functions may be defined as the function which represents only maximal or technically efficient production (38:174). Therefore, no unit can be more than 100% efficient. If we define a production function as the efficient input to output relationship, OLS can not be used to estimate the production function (2:268-269). Only frontier regression, stochastic frontiers, or other frontier estimation techniques

can be used to approximate this functional relationship. Using these types of models, any deviation from the efficient frontier is assumed to be due to inefficiency or stochastic error (31:305-306).

We are interested in approximating this functional relationship for use in a simulation in order to select the most accurate DEA envelopment frontier. No matter what technique is used to estimate the production function used in the simulation, we are primarily interested in approximating the input to output variable relationship using relevant variables which accurately estimate returns-to-scale.

Due to the DEA assumptions of Pareto-optimality and concave, monotonic functional forms, any variable with an OLS estimated coefficient which is negative should be closely examined before it is included in the simulation or DEA model. Theoretically, any increase in resource input should not decrease production. At the very worst, production should remain the same--otherwise the function would not be monotonic. If an inverse relationship between an input variable and an output variable is suggested by relevant economic theory or expert opinion, the analyst can rescale the input data. For example, theory may prescribe that when evaluating the relative efficiency of insurance salesmen, an increase in competing salesmen in a certain area actually reduces production of insurance sales. The variable representing number of competing insurance salesmen may be modeled by taking the inverse of the actual number of competitive salesmen. This guarantees the production relationship for competing salesmen will be monotonic. Presence of a negative OLS estimated coefficient for an increasing resource may also indicate collinearity of the input data (23:273).

The third phase of the PCA/ OLS/ Monte-Carlo Simulation strategy involves using the approximate OLS or frontier regression estimated production function in a Monte-Carlo simulation. The purpose of the simulation is to identify the most accurate envelopment frontier for use in the DEA model formulation. The estimated function in the simulation can be thought of as a relatively close approximation to the true but unknown functional relationships between the observed resource inputs and actual produced outputs. The DEA model formulation which most accurately estimates the efficiency of the simulated DMUs producing according to this production function should also be the most accurate DEA model formulation to estimate the efficiency of the actual DMUs.

Economic theory or past experience may prescribe the most appropriate form of the DEA envelopment frontier--i.e., Multiplicative to represent known, VRS, Cobb-Douglas production or CCR for known, CRS production. For instance, it would make intuitive sense to use a Multiplicative DEA model to determine efficient units in the natural gas pipeline industry because this specific industry demonstrates increasing marginal productivity (15:44). However, we can use the simulation model to evaluate any envelopment frontier-- BCC, CCR, Multiplicative with or without intercept, or Additive-- or different virtual multiplier linked cone constraints for a specific envelopment frontier.

The simulation model should represent the real world system as accurately as possible. The DMU sample size for each iteration should approximate the actual number of DMUs under evaluation. Similarly, the amount of resource input for each variable should approximate the actual resource usage for the DMUs under evaluation. The random variables used in the simulation to represent the inputs or resources should accurately

reflect not only the marginal distributions for each input, but also any joint distribution--or covariance--between the inputs (36:504). Simulation studies of known production functions indicate the level of covariance between input variables affects the ability of DEA to discriminate between efficient and inefficient DMUs. If a DEA model is misspecified and a highly correlated, relevant variable is excluded, the remaining variables in the DEA model will still contain some information about the excluded variable due to the high correlation (46:246-247).

Similarly, the distribution of the simulated efficiency scores should replicate the distribution of the actual efficiency scores. However, the analyst does not know the number of efficient DMUs or distribution of the true efficiency scores. The average number of efficient DMUs may be approximated using ratio efficiency analysis, expert opinion, or other prior efficiency evaluations. Specifying 25% of the DMU population as efficient is consistent with many empirical DEA efficiency studies (8:4). Similarly, based on historic DEA results, many analysts recommend using exponential or half normal distributions to model DMU efficiency scores (8:13).

There are a large number of MOEs which may be used to evaluate and eventually select the appropriate form of the envelopment frontier. Three MOEs to evaluate the performance of DEA models are:

- Minimize normalized Mean Absolute Deviation (MAD) of DEA estimated efficiency from actual efficiency
- Maximize the correlation coefficient between DEA estimated efficiency and actual efficiency

- Minimize the total number or average percentage of incorrectly identified DMUs

The normalized MAD from the actual efficiency scores is used in this research because different DEA model formulations use different metrics to measure the distance to the efficient frontier (16:34). While input oriented BCC and CCR formulations measure efficiency on a scale from 0 to 1, the Additive and Multiplicative models measure efficiency on a scale from $-\infty$ to 1. Normalizing these scores provides a more precise measure of the accuracy of each DEA model formulation when comparing the efficiency scores of different models.

Since DEA is a descriptive model and any calculated efficiency scores are relative measures based on the empirical envelopment frontier, correctly identifying efficient and inefficient DMUs may well be the most important measure for any DEA model. As already mentioned, a descriptive statistic which measures this accuracy is the Average Percent Error Rate (APER).

Based on the Monte-Carlo simulation results, the analyst selects the most accurate and appropriate DEA envelopment frontier. Combined with the results of the PCA/OLS analysis used to select the relevant input variables, the analyst can then construct an accurate DEA model formulation. The efficiency information obtained from this DEA model can be subsequently used in further analysis, including OLS estimation of a production function (8:8-10). This production function should be improved by the additional information the DEA analysis provides, and may be used for parameter estimation or forecasting. The amount of improvement the DEA information provides

may be objectively measured by the OLS model's increased adjusted R^2 and decreased forecast MAPEs both prior to and after including the DEA efficiency information.

IV. Results of the DEA Modeling Strategy

4.1 Identifying Relevant DEA Input Variables

The first stage of the DEA modeling strategy uses PCA and OLS to identify relevant variables to be used in the DEA model. Again, relevant variables are defined as "...somewhat related experientially, statistically, and/or conceptually to the production process." (16:427). Using PCA on the eight recruiting resource variables used in the current FAARR DEA model for 3rd Quarter FY96 thru 2nd Quarter FY97, the author identified four underlying dimensions of the recruiting resource data. The factor loadings matrix, associated eigenvalues, and scree test are depicted below in Table 4.1 and Figure 4.1. A variable is considered to heavily load on a factor if the absolute value of its score in the factor loadings matrix is greater than 0.5. Since the DoD Recruiter variable does not heavily load on any factor, it may not be a relevant variable for the DEA model.

Table 4.1: Recruiting Data PCA Factor Loadings Matrix

	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6	Factor 7	Factor 8
Eigenvalue	0.55	0.695	0.381	0.89	0.143	1.23	1.605	2.505
OPR	0.501	7.22E-03	-0.103	0.066	0.179	0.113	-0.058	-0.828
TVGRP	0.237	-0.026	0.359	-0.104	-0.087	-0.09	0.88	-0.116
RADGRP	-0.036	-0.338	-0.322	-0.222	4.16E-03	0.565	0.608	0.207
MAGGRP	0.02	0.461	0.161	-0.245	-7.19E-03	0.805	-0.207	0.101
LOCAL\$	-0.022	0.058	0.107	0.061	0.29	-0.089	0.234	0.913
DODREC	-0.222	0.432	-0.178	-0.448	0.061	-0.387	0.391	-0.475
POP	-0.194	0.264	-0.07	0.744	-1.68E-03	0.209	0.45	-0.298
UNEMP	-0.392	-0.33	0.253	-0.086	0.126	0.201	-0.068	-0.779

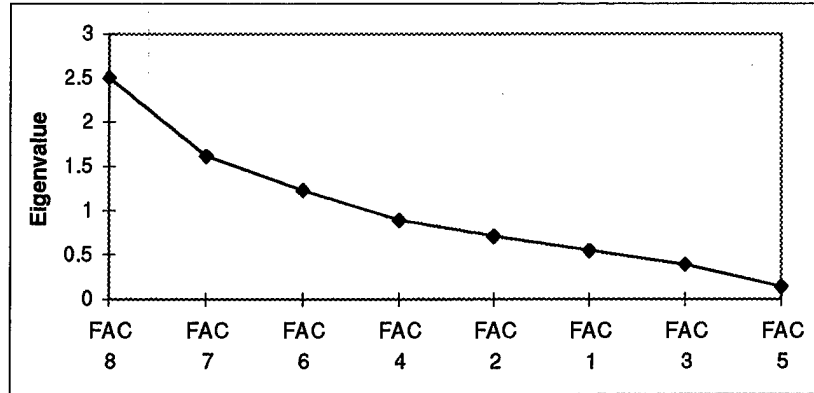


Figure 4.1: Recruiting Data PCA Eigenvalue Scree Test

Using an OLS model to test the statistical significance of the time series and cross sectional data for the eight recruiting resource variables resulted in a similar conclusion. A linear OLS model was used with indicator variables to adjust for seasonality. The DoD recruiter variable was only significant at the .97 level and the Television GRP variable was only significant at the .839 level (Table 4.2). Again, statistical evidence supports dropping these possibly irrelevant variables from the DEA model.

Table 4.2: OLS Statistical Significance Results for Recruiting Resource Variables

Variable	Coefficient	P-Value
Intercept	-1.036	6.88E-01
OPR	1.057	1.00E-16
TVGRP	-0.00988	0.839
RADGRP	-0.201	3.36E-06
MAGGRP	0.256	0.0002
LOCAL\$	0.0293	0.206
DODREC	-0.00937	0.973
POP	0.0673	0.285
UNEMP	0.143	0.053
Quarter 2	-0.106	0.125
Quarter 3	-0.161	0.0008
Quarter 4	-0.0893	0.0458

Additionally, including the DoD recruiter in any OLS model or DEA model may result in problems with multicollinearity. Several diagnostic tests indicate the original eight

variable data set is highly correlated. First, as the correlation matrix in Table 4.3 indicates, the DoD Recruiter variable is highly negatively correlated with the dependent variable (GSMA contracts) and the OPR and population independent variables. However, the DoD Recruiter variable is not statistically significant in the OLS regression model (23:273). Second, the sum of the inverses of the eigenvalues of the correlation matrix equal 15.84. If the independent variables were all orthogonal--not correlated--this statistic would equal eight. Large values for this statistic indicate severe collinearity (23:274). Finally, the Variance Inflation Factors (VIF) for each of the independent variables are depicted in Table 4.4. Again, due to the large value of this statistic for the DoD recruiter variable, the author suspects multicollinearity problems from the data set (40:658).

The results of the PCA and OLS analysis and the probable problem with multicollinearity indicate the DoD recruiter variable should not be included in the DEA model.

Table 4.3: Recruiting Data Correlation Matrix

	GSMA	OPR	TVGRP	RADGRP	MAGGRP	LOCAL\$	DODREC	POP	UNEMP
GSMA	1								
OPR	0.74669	1							
TVGRP	0.151371	0.092595	1						
RADGRP	-0.1979	-0.00202	-0.18847	1					
MAGGRP	0.014425	-0.14472	0.368259	0.19681	1				
LOCAL\$	0.151672	0.205082	0.256247	-0.03095	0.105267	1			
DODREC	-0.59228	-0.74554	0.109454	-0.0007	0.215668	-0.12813	1		
POP	0.42432	0.462357	0.016106	-0.00236	-0.02502	0.149906	-0.69622	1	
UNEMP	0.296014	0.21709	0.346822	-0.16563	-0.0597	0.072636	-0.30603	0.210493	1

Table 4.4: Recruiting Resource Data Variance Inflation Factors

Variable	VIF
DODREC	4.3426
OPR	2.5930
POP	2.0779
TVGRP	1.7604
UNEMP	1.3979
MAGGRP	1.3756
RADGRP	1.1448
LOCAL\$	1.1222
Mean VIF	1.9768

The results of the statistical and qualitative analysis of the input variables indicate only five variables should be used in the DEA model-- recruiters, print GRPs, local advertising expenditures, population, and the unemployment rate. The TV GRP variable was not used because of lack of statistical significance in the OLS screening. The Radio GRP variable model was not used because of the negative coefficient in both the OLS screening and correlation matrix. The DoD Recruiter variable was not used due to lack of statistical significance in both the PCA analysis and OLS analysis and its high correlation to other input variables. Table 4.5 summarizes the analysis and screening of the relevant input variables based upon the variables' accuracy; intercorrelation; and statistical, experiential, and conceptual relation to GSMA contracts.

Table 4.5: Recruiting Resource Variable Analysis and Screening

Input Variable	Accuracy	Inter-Correlation	Statistically Related			Positive OLS/ Correlation Coefficient	Experientially/ Conceptually Related	Used in DEA Model
			Correlation Coefficient	Factor Analysis	OLS Screening			
Recruiters	High	Medium	High	Yes	High	YES	High	YES
Print GRPs	Low	Low	Low	Yes	High	YES	Medium	YES
Local Advertising	High	Low	Low	Yes	Medium	YES	Medium	YES
Population	Medium	Medium	Medium	Yes	Medium	YES	High	YES
Unemployment Rate	High	Low	Medium	Yes	High	YES	Medium	YES
Television GRPs	Low	Low	Low	Yes	Low	NO	Medium	NO
Radio GRPs	Low	Low	Low	Yes	High	NO	High	NO
DoD Recruiters	Medium	High	Medium	No	Low	NO	High	NO

4.2 Estimating a Recruiting Production Function

The second phase of the DEA model formulation strategy is to estimate an approximate production function using the relevant variables identified in the PCA/OLS analysis. Two production functions were estimated for the recruiting data using both OLS and a deterministic frontier estimation technique known as efficient frontier benchmarking (31: 306).

To estimate the OLS model for the simulation, step-wise linear regression was used with the five relevant recruiting resource variables identified in the previous section and “dummy” indicator variables to account for seasonality. The 2nd, 3rd, and 4th quarters of each fiscal year were represented by variables named QTR2, QTR3, and QTR4. Pooled, times series and cross sectional variables for the four quarters from 3rd QTR FY96 thru 2nd QTR FY97 (33:396) were used in a OLS model with a Cobb-Douglas, log-log formulation. The dependent variable was the number of GSMA contracts. The general mathematical formulation of the production function was:

$$y_j = \alpha_0 \prod x_{ij}^{\beta_i}$$

where

y_j \equiv output for battalion j

x_{ij} \equiv input i for battalion j

β_i \equiv coefficient for input i

α_0 \equiv intercept term

This function was approximated by OLS using the natural logarithms of the independent and dependent variables.

Using both forward and backward step-wise regression and initial OLS production function model was:

$$\ln(\text{GSMA}) = -1.842005608 + 0.991010447*\ln(\text{OPR}) + 0.119530613*\ln(\text{MAGGRP}) + \\ 0.025509477*\ln(\text{LOCAL\$}) + 0.096590765*\ln(\text{POP}) + 0.176151048*\ln(\text{UNEMP}) - \\ 0.11413 (\text{QTR3})$$

This model's adjusted R^2 was 60.06%, and variables had positive coefficients and were significant at the .25 level. The local advertising expenditure variable was statistically significant at only the .25 level. Although this variable may not have been included in a more rigorous OLS model, expert opinion suggests that it is a critical resource in the recruiting process. Variable QTR2 and QTR4 were not significant at the .25 level and were not used in the final OLS model. Residual analysis indicated no problems with seasonality or a trend in the residuals. Durbin-Watson statistics were calculated for the three different OLS models for successive four quarter time periods. The Durbin-Watson statistics-- 2.150, 2.238, and 2.255, respectively-- did not indicate the error terms were correlated. Figure 4.2 illustrates the fit of the 161 observations of actual GSMA contract production to the OLS estimated production function model for the 41 battalions for the four quarters. Because OLS was used to fit the production function, actual contract production may be less than or greater than estimated contract production.

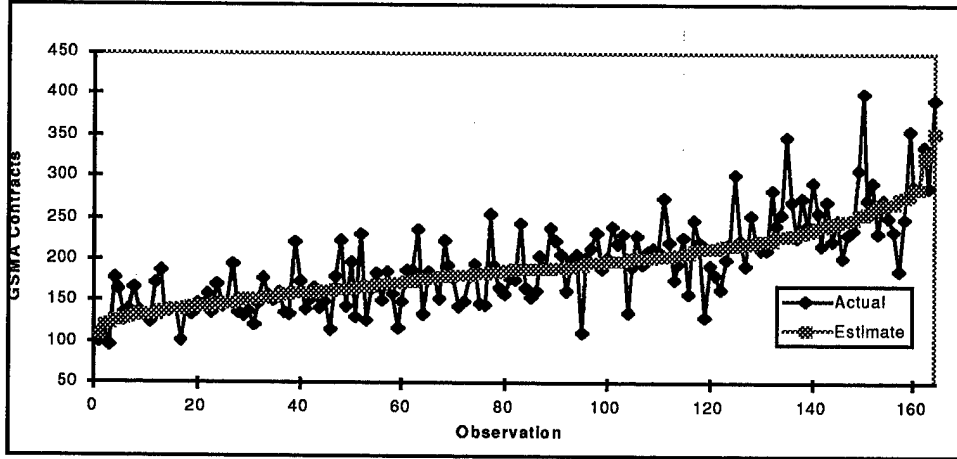


Figure 4.2: Actual GSMA Contract Production versus OLS Estimate

The second estimated production function is a frontier production function. The efficient frontier benchmarking model used is a straight forward, deterministic, frontier estimation model which minimizes the sum of the deviations from the frontier across all DMUs (31:306). Any deviation from the efficient frontier is assumed to be due to inefficiency. No assumption is made concerning the existence or distribution of an error term. The mathematical programming formulation for this model is:

$$\text{Minimize: } \sum_j \epsilon_j \quad (10)$$

subject to

$$\ln(y_j) = \alpha_o + \sum_i \beta_i \ln(x_{ij}) + \partial_k D_k - \epsilon_j, \text{ for all } j=1, \dots, n \quad k=1, \dots, 3 \quad (11)$$

$$\epsilon_j \geq 0, \text{ for all } j=1, \dots, n \quad (12)$$

$$\beta_i \geq 0, \text{ for all } i \quad (13)$$

where

ϵ_j \equiv deviation from the efficient frontier for battalion j

y_j \equiv output for battalion j

x_{ij} \equiv input i for battalion j

β_i \equiv coefficient for input i

α_o \equiv intercept term

∂_k \equiv coefficient for indicator ("dummy") variable for season k

D_k \equiv a 0 or 1 indicator ("dummy") variable for season k

The use of the indicator-- or “dummy”-- variables in equation (11) of the frontier model allow for the inherent seasonality of the recruiting process. If these indicator variables were not included, any deviation from the frontier due to seasonality would be attributed to a battalion’s inefficiency. Both the Kurskal-Wallis non-parametric tests and practical experience support the assumption that the recruiting process is inherently seasonal. Many new recruits enter active service following graduation from high school in the May-June time frame. Equation (13) ensures the function will be monotonic with regard to the utilization of any recruiting resource.

As with the OLS estimation of the recruiting production function, the same five pooled, time-series and cross sectional variables were used: recruiters, print GRPs, local advertising expenditures, population, and the local unemployment rate. Additionally, variables QTR2, QTR3, and QTR4 were included to account for seasonality. The resulting efficient frontier benchmarking production function was:

$$\begin{aligned} \ln(\text{GSMA}) = & - 1.953978 + 0.968927*\ln(\text{OPR}) + 0.00*\ln(\text{MAGGRP}) + \\ & 0.017484*\ln(\text{LOCAL\$}) + 0.183641*\ln(\text{POP}) + 0.352645*\ln(\text{UNEMP}) \\ & - 0.006562*(\text{QTR2}) - 0.047544*(\text{QTR3}) + 0.099867*(\text{QTR4}) \end{aligned}$$

The estimated coefficients for this function did not radically differ from the OLS model. However, it should be noted that the estimated coefficient for print GRPs was zero while the print GRP coefficient was .11953 using the OLS estimated production function.

Figure 4.3 below graphically illustrates the fit of this efficient frontier model to the actual data. Because a frontier estimation technique was used to fit the production function, actual contract production must be less than or equal to the estimated contract production.

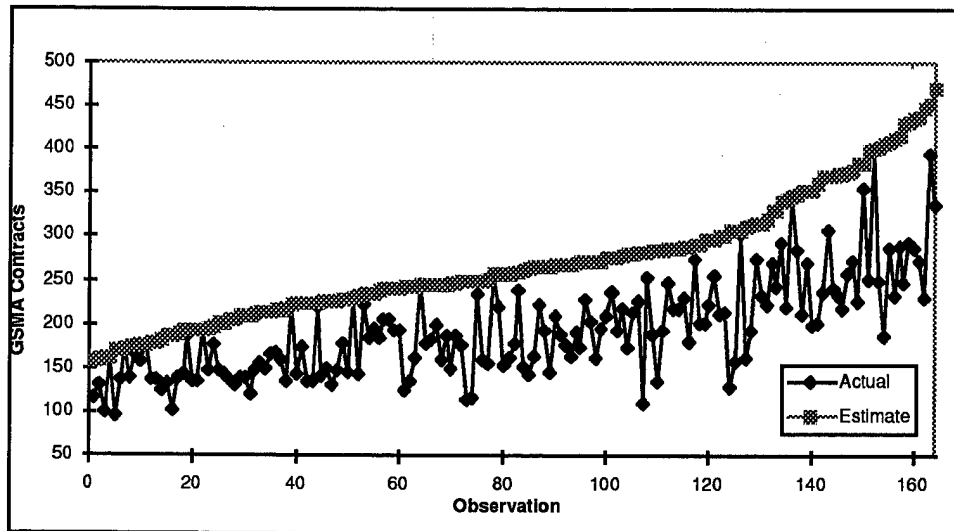


Figure 4.3: Actual GSMA Contract Production versus Efficient Frontier Benchmarking Estimate

4.3 Simulation of Recruiting Production Function

The third phase of PCA/ OLS/ Monte-Carlo Simulation methodology involves using the OLS estimated production function or efficient frontier production function in a Monte-Carlo simulation to identify the most accurate DEA envelopment frontier. Using GAMS software, four simulations were constructed using different known production functions-- the OLS estimated production function with and without an error term and the frontier production function with and without an error term. These four production functions were used to evaluate the accuracy of the Additive, output oriented BCC, output oriented CCR, Multiplicative, and Multiplicative without intercept DEA models. The output oriented BCC and CCR models were chosen because the random technical efficiency term, η_j , was applied to the simulated production function output, and the efficiency scores of the output oriented DEA models would be more consistent estimators of a DMUs true efficiency (5:240).

Each simulation replication generated 41 simulated recruiting battalions (DMUs) using a five dimensional random input vector of recruiting resources selected from a multi-variate normal distribution estimated from 1st Quarter FY97 recruiting data. The five random variables-- recruiters, print GRPs, local advertising expenditures, population, and the local unemployment rate-- represented the relevant recruiting resources identified by the PCA and OLS analysis. These resource inputs were used in the OLS or frontier production functions to calculate a recruiting battalion's efficient production. The efficient contract production for recruiting battalion j (DMU j) was multiplied by a random variable, η_j , representing the actual or "true" technical efficiency of simulated recruiting battalion j , where $\eta_j \in [0,1]$. The product of the efficient production and the random technical efficiency score is the simulated recruiting battalion's actual production observed by the DEA model. Similar to the simulation used to validate the FAARR model, the random, actual efficiency scores were selected from a truncated normal distribution estimated using the results from the FAARR DEA model. For this specific distribution, approximately 11.2% of the recruiting battalions are efficient. Although historical DEA efficiency studies have concluded approximately 25% of DMUs are efficient (8:4) and are usually distributed exponentially or from a half-normal distribution (8:13), the truncated normal distribution used in this simulation was the result of past DEA analysis of actual Army recruiting battalions. It was judged that the results of past DEA modeling for this specific data would be a more accurate representation of actual recruiting efficiency than general results from across the DEA literature.

For the two simulations with an error term, a normally distributed random variable with a mean of zero and a standard deviation of ten was added to the calculated production of GSMA contracts. Given the average contract production for actual recruiting battalions, 95% of the random errors should be within +/- 10% of a simulated recruiting battalion's efficient production. One hundred simulation replications were conducted for each model.

The mathematical formulations of the four simulation production functions (OLS estimated and frontier production functions both with and without an error term) were:

$$y_j = \alpha_0 \prod x_{ij}^{\beta_i} \eta_j + \epsilon_j$$

where

- y_j \equiv output of theoretical battalion j
- α_0 \equiv estimated production function intercept term
- x_{ij} \equiv random input i for battalion j
- β_i \equiv estimated production function coefficient for input i
- η_j \equiv technical efficiency for battalion j
- ϵ_j \equiv random error term for battalion j (if applicable)

Both the OLS estimated and frontier benchmarking production functions were Increasing Returns-to-Scale (IRS) functions since the sum of the estimated input coefficients was greater than 1.

Three MOEs were used to evaluate each simulation model: the normalized Mean Absolute Deviation (MAD) of the DEA estimated efficiency from actual efficiency, the correlation coefficient between estimated and actual efficiencies, and the average percentage of incorrectly identified battalions (APER). The APER consists of inefficient battalions which were classified as efficient and efficient battalions which were classified as inefficient. A DEA envelopment with a lower normalized MAD and a smaller APER is a

more accurate model. Similarly, a DEA envelopment with a higher correlation coefficient is a more accurate model.

4.4 Simulation Results

The results of the four simulations and the average results across all four simulations are depicted in the five tables below. The first column of each table indicates which DEA model was evaluated. The second column indicates the normalized Mean Absolute Deviation (MAD) of the DEA estimated efficiency from actual efficiency. The third column indicates the correlation coefficient between estimated and actual efficiencies. The fourth and fifth columns indicate the rates at which inefficient battalions were incorrectly classified as efficient and efficient battalions were incorrectly classified as inefficient. The sixth column indicates the total APER classifying the simulated battalions.

Table 4.6: Simulation Results for OLS Production Function without Error Term

DEA Model	Correlation Coefficient	Normalized MAD	% Incorrect Classification of DMUs		
			NotEFF EFF	EFF NotEFF	Total APER
Additive	0.4894	0.0118	0.00	51.13	45.44
BCC	0.6410	0.0195	1.30	48.06	42.90
CCR	0.6764	0.0116	36.02	22.44	24.24
Multiplicative	0.6099	0.0113	0.00	45.46	40.41
Multiplicative w/o Intcpt	0.6118	0.0127	20.24	22.41	22.54

Table 4.7: Simulation Results for Frontier Production Function without Error Term

DEA Model	Correlation Coefficient	Normalized MAD	% Incorrect Classification of DMUs		
			NotEFF EFF	EFF NotEFF	Total APER
Additive	0.4944	0.0111	1.12	50.32	44.90
BCC	0.6290	0.0192	1.61	47.89	42.83
CCR	0.5975	0.0126	41.31	21.40	24.07
Multiplicative	0.5777	0.0113	0.40	45.27	40.32
Multiplicative w/o Intcpt	0.5668	0.0127	27.36	22.37	23.29

Table 4.8: Simulation Results for OLS Production Function with Error Term

DEA Model	Correlation Coefficient	Normalized MAD	% Incorrect Classification of DMUs		
			NotEFF EFF	EFF NotEFF	Total APER
Additive	0.5144	0.0129	1.75	50.61	45.15
BCC	0.5420	0.0172	17.61	45.16	41.93
CCR	0.5998	0.0136	43.42	21.53	24.22
Multiplicative	0.4969	0.0146	16.09	42.49	39.51
Multiplicative w/o Intcpt	0.5360	0.0153	36.93	22.81	24.85

Table 4.9: Simulation Results for Frontier Production Function with Error Term

DEA Model	Correlation Coefficient	Normalized MAD	% Incorrect Classification of DMUs		
			NotEFF EFF	EFF NotEFF	Total APER
Additive	0.4433	0.0147	11.40	48.67	44.24
BCC	0.5715	0.0179	15.06	46.32	42.59
CCR	0.5777	0.0135	46.11	21.07	24.12
Multiplicative	0.5114	0.0143	14.07	43.51	40.02
Multiplicative w/o Intcpt	0.5336	0.0148	35.76	22.45	24.39

Table 4.10: Average Simulation Results For All Production Functions

DEA Model	Correlation Coefficient	Normalized MAD	% Incorrect Classification of DMUs		
			NotEFF EFF	EFF NotEFF	Total APER
Additive	0.4854	0.0126	3.57	50.18	44.93
BCC	0.5959	0.0185	8.90	46.86	42.56
CCR	0.6128	0.0128	41.72	21.61	24.16
Multiplicative	0.5490	0.0129	7.64	44.18	40.07
Multiplicative w/o Intcpt	0.5620	0.0139	30.07	22.51	23.77

Analysis of the simulation results yield some very interesting conclusions. The Additive DEA envelopment consistently performed the worst in terms of the correlation coefficient and APER in all four simulations. The BCC DEA envelopment consistently performed the worst in terms of the normalized MAD in all four simulations. The CCR envelopment performed the best in terms of the correlation coefficient in three of the four simulations. Using APER as the evaluation criteria, the Multiplicative DEA envelopment without an intercept term performed the best when the production function had no error term. The CCR DEA envelopment had the best APER when an error term was included in the simulation. Compared to the CRS DEA models, all of the VRS DEA models-- Additive, BCC, Multiplicative-- had significantly higher rates incorrectly classifying

inefficient battalions as efficient. For all four simulation, VRS envelopments incorrectly classified inefficient DMUs as efficient at almost twice the rate of the CRS models. Alternately, the CRS envelopments incorrectly classified efficient DMUs as inefficient at a much higher rate than the VRS envelopments.

Analysis of these results indicates a CRS model-- either CCR or Multiplicative envelopment without an intercept term-- are the most accurate envelopment shapes when the production process is IRS. Using a VRS DEA envelopment-- BCC, Multiplicative, or Additive-- to estimate the efficiency of DMUs producing output according to an IRS process results in upwardly biased efficiency scores. VRS models attribute a DMU's less than efficient production to a change in the production processes returns-to-scale and not to the DMU's actual inefficiency. This may indicate that selecting the appropriate shape of the DEA envelopment is the most important step in the DEA modeling process.

Using the three MOEs, the Additive and BCC envelopments are clearly the least accurate models for this simulated production function. The CCR envelopment consistently outperformed both Multiplicative envelopments in regards to the correlation coefficient and consistently outperformed the Multiplicative envelopment without an intercept term in regards to the normalized MAD. The Multiplicative envelopment without an intercept term averaged only .4% lower APER than the CCR envelopment across all four simulations. The author judges the CCR envelopment to be the most accurate DEA formulation given its superiority in terms of two MOEs and relatively high accuracy in correctly classifying individual DMUs. The Multiplicative envelopment without an intercept term is the second best choice. The two best performing

envelopments, the CCR model and the Multiplicative model without an intercept term, were both CRS. Table 4.11 summarizes the results of the four simulations.

Table 4.11: Simulation Result Summary

Simulation	Measures of Effectiveness		
	Correlation Coefficient	Normalized MAD	TOTAL APER
OLS w/o Error Term			
Best Performing	CCR	Multiplicative	Mult w/o Intcpt
Worst Performing	Additive	BCC	Additive
Frontier Function w/o Error Term			
Best Performing	BCC	Additive	Mult w/o Intcpt
Worst Performing	Additive	BCC	Additive
OLS with Error Term			
Best Performing	CCR	Additive	CCR
Worst Performing	Multiplicative	BCC	Additive
Frontier Function with Error Term			
Best Performing	CCR	CCR	CCR
Worst Performing	Additive	BCC	Additive
Average across four simulations			
Best Performing	CCR	Additive	Mult w/o Intcpt
Worst Performing	Additive	BCC	Additive

Other DEA simulation studies of known production functions support these conclusions for IRS functions. Banker, Chang, and Cooper's (5) simulation study determined the estimated efficiency scores from a CCR envelopment had a lower MAD from the true efficiencies than a BCC envelopment using a simulated two input, one output, IRS Cobb-Douglas production function with a sample size of 50 DMUs. The BCC envelopment had a lower MAD than the CCR envelopment when the function had DRS (5:238-239).

Smith's simulation study (46) is more comprehensive because he used a known Cobb-Douglas functional form and varied the number of input variables from two to six and varied the DMU sample size from 10 to 80 DMUs. His research was focused on the affects of DEA model misspecification and he primarily evaluated the output-oriented

CCR model (46:236). However, while researching the affects on varying the DEA model's assumption concerning returns to scale, he compared both the CCR and BCC models. Smith concluded for a known CRS process with five resource inputs and a sample size of 40 DMUs, VRS BCC model efficiency scores were 11.7% higher than CRS CCR model scores (46:245). Smith concluded using the BCC model to evaluate efficiency of DMUs results in an increase in estimated efficiency (46:244). Banker, Chang, and Cooper reach similar conclusions concerning the choice of the envelopment frontier (8:239).

The results of this research are similar. The BCC model consistently overestimated the average DMU efficiency using all four simulated IRS production functions. On average, the BCC envelopment overestimated actual DMU efficiency scores by 6.25%. As Table 4.10 depicts, the BCC model incorrectly classified inefficient DMUs as efficient, therefore overestimating their actual efficiency score, 46.86% of the time. Again, it appears that an incorrect choice of a VRS DEA envelopment frontier for a CRS (46:245) or IRS (this research) function results in upwardly biased efficiency estimates. In an attempt to maximize each DMU's efficiency score, the model attributes a DMU's less than efficient production to a change in the production function's returns-to-scale and not to any inherent DMU inefficiency.

In summary, this analysis indicates a CCR DEA model using OPR, MAGGRP, LOCAL\$, POP, and UNEMP as variables is the most accurate DEA model to estimate the efficiency of U.S. Army recruiting battalions.

4.5 Efficiency Estimates Using CCR Model Formulation

Using the five variable CCR model identified in the previous section, all 41 Army recruiting battalions were evaluated for seven consecutive quarters. Nineteen battalions were consistently rated inefficient for all seven quarters. Only one battalion, battalion 6G, was rated efficient for all seven quarters. Battalion 1B was rated efficient for five quarters. In comparison, using the original eight variable, FAARR DEA model, twenty-five battalions were consistently rated inefficient for all seven quarters. No battalions were rated efficient for all seven quarters, but battalion 3T was rated efficient for six quarters. Table 4.12 compares the estimated efficiency scores and percentage of efficient DMUs for the five variable CCR model and the eight variable FAARR DEA model. As this research and the referenced literature indicate, for a particular set of DMUs, the specification of the DEA model can result in drastically different efficiency scores and number of efficient DMUs. If DEA efficiency information is to be useful as a management tool to evaluate DMU performance, we must have some confidence or objective measure of a DEA model's accuracy. Without *a-priori* knowledge of the production process, the analyst has little way of knowing which DEA model will yield the most accurate estimate of actual DMU efficiency.

Table 4.12: Comparison of DEA Model Efficiency Scores

DEA Model	Average % DMUs Rated Efficient	Average Efficiency	STD DEV Efficiency
5 variable CCR	15.68	0.8138	0.2345
8 variable FAARR	10.1	0.8739	0.1099

In summary, inferences concerning DEA model misspecification (46) are:

- Omitting a relevant variable may reduce estimated efficiency
- Including an irrelevant variable may increase estimated efficiency
- Omitting a highly correlated, relevant variable is not as serious as omitting a non-correlated relevant variable
- Including an irrelevant variable is not as serious as omitting a relevant variable
- Assuming variable returns to scale when the process is CRS or IRS may increase estimated efficiency
- DEA models with fewer input variables and smaller numbers of DMUs may be more sensitive to invalid assumptions concerning the production process returns-to-scale than to an invalid choice of input variables

These inferences directly correlate to the following DEA model building tactics:

- Err on the side of including irrelevant variables rather than excluding relevant variables
- Gather as much information as possible to accurately determine the classification of the production process's returns-to-scale⁴

As the results of the FAARR evaluation of a simulated CRS production function demonstrate, assuming a VRS DEA model for a CRS production process leads to increased, erroneous estimates of efficiency.

4.6 Practical Application of DEA Efficiency Information in Combined OLS/DEA Models

⁴ OLS regression techniques, DEA Most Productive Scale Size estimates (6:34-35), translog functions (5:236-237), economic theory (38:235-238), and expert opinion (15:44) are all useful in determining if a production process is DRS, IRS, CRS, or VRS.

Using an accurate DEA model, the analyst can categorize DMUs as efficient or inefficient. The following example illustrates how this information can be used in a conventional OLS model to both estimate parameters and forecast future production. Although the OLS model provides a relatively good fit and more accurate forecasts than the FAARR model, the OLS model is general in nature and is only intended to illustrate the use of DEA efficiency information. The author hypothesizes that more accurate DEA efficiency information will improve an OLS model in general, and the USAREC model in particular (8:2).

Stepwise regression was used to estimate a causal, Cobb-Douglas OLS model using four quarters of pooled, time series and cross sectional recruiting data. The goal was to develop a single model, using the same variables, which would accurately fit three separate sets of data-- a rolling horizon of four quarters of recruiting data. This single model was used to estimate the responses for 1st QTR FY97 production, 2nd QTR FY97 production, and 3rd QTR FY97⁵ production.

The initial independent variables considered for the model included the eight recruiting resource variables, three indicator variables for seasonality (QTR2, QTR3, and QTR4), and four indicator variables to identify specific recruiting brigades (BDE2, BDE3, BDE4, and BDE5). Since recruiting brigades are organized geographically, these indicators variables account for geographic as well as organizational variations. The dependent

⁵ Battalion 5C was excluded from the forecast for 3rd QTR FY97. In a one time administrative measure, the new battalion commander wrote off 23 GSMA contracts for 3rd QTR FY97. The battalion commander questioned the ability of these enlistees to complete their time in the Delayed Entry Program (DEP) prior to entering active duty. This administrative accounting measure significantly biased the forecast accuracy MOEs for the 3rd QTR FY97 forecast.

variable was the number of GSMA contracts. The best fitting, step-wise regression model for all three sets of data was:

$$\ln(\text{GSMA}) = \alpha_0 + \beta_1 \ln(\text{OPR}) + \beta_2 \ln(\text{MAGGRP}) + \beta_3 \ln(\text{LOCAL\$}) + \beta_4 \ln(\text{POP}) + \beta_5 \ln(\text{RADGRP}) + \delta_1(\text{QTR3}) + \delta_2(\text{BDE 5})$$

Note that the variables included in this OLS estimated model differed from the variables in the OLS estimated simulation production function use to select the most accurate DEA model. Variables which may only have been significant at the .25 level were included in the simulated production function OLS model, but were not included in this more rigorous forecasting OLS model.

The average adjusted R^2 for this model across the three time periods was 70.4%. All variables except the print GRP were significant at the .05 level. The Print GRP variable was significant at the .15 level. This causal model represented the OLS estimate of GSMA contracts at time t:

$$y_t = \alpha_0 \prod x_{i(t)}^{\beta_i}$$

To predict future contracts the author used the estimated parameter coefficients from time period t with the recruiting resources for time t+1:

$$y_{t+1} = \alpha_0 \prod x_{i(t+1)}^{\beta_i}$$

For the forecasts, the author used the recruiting resource levels, $x_{i(t+1)}$, the actual levels at time period t+1 for recruiters, print GRPs, radio GRPs, and local advertising expenditures. These variables are all discretionary and controllable by USAREC. The author used the actual population at time t as the population estimate at time t+1. This variable is not

discretionary and the best available estimate for the population at time period t for time period t+1 was the population at time period t.

As Table 4.13 indicates, this model produced forecasts for the three quarters with an overall model MAPE of 4.06%, an average MAPE per recruiting battalion (DMU) of 15.11%, and an average maximum MAPE across all battalions of 50.17%. The average battalion MAPE is the most accurate indicator of forecast accuracy. The overall model MAPE is merely the sum of the individual battalion forecasts compared to the actual USAREC wide contract production. The overall model MAPE statistic contains no information about the forecast accuracy for individual battalions. A particular model may have a low overall model MAPE although individual battalion forecasts deviate drastically from actual battalion production.

DEA efficiency information was then used to estimate a similar times-series, cross sectional, OLS model which included an additional “dummy” indicator variable (defined as variable DEA) representing DEA efficient battalions:

$$\ln(\text{GSMA}) = \alpha_0 + \beta_1 \ln(\text{OPR}) + \beta_2 \ln(\text{MAGGRP}) + \beta_3 \ln(\text{LOCAL\$}) + \beta_4 \ln(\text{POP}) + \beta_5 \ln(\text{RADGRP}) + \delta_1(\text{QTR3}) + \delta_2(\text{BDE 5}) + \delta_3(\text{DEA})$$

where

δ_3 = an indicator (“dummy”) variable equal to 1 if the recruiting battalion is rated efficient and equal to 0 if the recruiting battalion is rated inefficient

The indicator variable for DEA efficiency (DEA) affects only the intercept of the function for the DEA efficient units. Indicator variables for recruiting resources, which would affect the slope of the function for DEA efficient units, were not significant at the .10 level and were not used. This step-wise OLS model also initially included five indicator

variables for regions of the country which represented the five recruiting brigades. The 5th Brigade's region, represented by variable BDE5, was the only region which was statistically significant.

Since there was no accurate forecast for a recruiting battalion's efficiency at time period t to be used in the forecasts for time period $t+1$, all battalions were assumed to be inefficient. Data to calculate a recruiting battalion's efficiency at time t is not available until the beginning of time period $t+1$. This assumption resulted in more accurate, but slightly downward biased, forecasts.

Table 4.13 depicts the average results of the three forecasts for the first three quarters of FY97 using various forecasting models. For comparison, naive and four quarter moving average forecast results are included in addition to the OLS forecasts with and without DEA efficiency information from the various DEA envelopments. OLS models using the DEA efficiency information as intercept indicator variables are referred to as OLS/DEA models.

Table 4.13: Comparison of Average Forecast Results for OLS and OLS/DEA Models

Average: 1st QTR FY97 thru 3rd QTR FY97				
	MAPE	Ave. BN MAPE	Max. BN MAPE	Adj R ²
Naive Forecast 1	9.42	15.61	54.80	N/A
Four Quarter Moving Average	6.67	10.89	38.14	N/A
Cobb-Douglas OLS	7.47	15.11	50.17	70.40
OLS/DEA Models (DEA efficiency indicator variable affects intercept)				
FAARR DEA* (VRS)	7.05	14.83	50.86	70.82
Additive (VRS)	10.20	15.87	50.85	75.91
BCC (VRS)	9.66	15.50	50.23	75.48
CCR (CRS)	6.55	14.65	48.24	75.84
Multiplicative (VRS)	9.41	15.43	51.05	75.08
Multiplicative w/o Intcpt (CRS)	6.73	14.73	48.57	73.27
A DMU is classified as efficient if its efficiency Score >.999				
*Original 8 Variable FAARR DEA Model Formulation				

Analysis of the regressions using OLS and combined OLS/DEA models suggests the DEA efficiency information available from all DEA models improved the fit of the regression as measured by the adjusted R^2 . Efficiency information from certain DEA models also improved the average accuracy of the forecasts. It is interesting to note, of the five classical DEA models, the two OLS/DEA models using CRS envelopments (CCR and Multiplicative without intercept) provided more accurate forecasts than the OLS model. The three OLS/DEA models using the VRS DEA envelopments-- Additive, BCC, and Multiplicative-- all had higher average battalion MAPEs than the OLS model. Since CRS DEA models seem to provide more accurate efficiency information, this evidence supports the assumption the recruiting production process is CRS.

Although not entirely conclusive, the regression and forecast results from the combined OLS/DEA models do support the previous conclusions regarding the most accurate DEA model formulation. If a specific DEA envelopment is more accurate estimating true DMU efficiencies, we would expect, everything else being equal, the combined OLS/DEA model using the more accurate DEA indicator variables to have better forecasts and a higher adjusted R^2 . The conclusions of our previous analysis indicated the CCR envelopment was the most accurate DEA model followed by the Multiplicative envelopment without an intercept term. The BCC and Additive DEA models were the least accurate. The regression results indicate the CCR model provides the most accurate forecasts, but the Additive model has the highest adjusted R^2 , followed closely by the CCR model. The high adjusted R^2 for the Additive model is contradictory to what we would expect but may not be significant. The CCR OLS/DEA model's adjusted R^2 was only .07 less than the

Additive OLS/DEA model's adjusted R^2 . Overall, the regression and forecast results support the conclusion the five variable CCR DEA model is probably the most accurate DEA model.

Analysis of alternate models' forecast accuracy also indicates that simple time-series models may be more accurate than causal OLS models. Similar to other times-series models, the four quarter moving average forecast model only relies on a battalion's past contract production to forecast future production. Time-series models make no assumption concerning the underlying production process or the use of resources and they are simple and very easy to construct. However, time-series models provide no estimates of resource parameters or output elasticities. The four quarter moving average forecast was provided for illustration purposes only, but this model not only has the lowest average battalion forecast MAPE, but it also has the smallest maximum battalion MAPE. The Naive Forecast 1 simply forecasts the upcoming quarters production using the actual production from the previous quarter. As stated in the third chapter, forecasts for the Naive Forecast 1 model were more accurate than the forecasts from the original FAARR model for all MOEs. This analysis of time-series forecasting models indicates a time-series model may provide USAREC more accurate forecasts than a causal OLS forecasting model.

V. Conclusions and Recommendations

5.1 Conclusion

The first part of this research evaluated the accuracy and robustness of the GSMA contract forecasts for the Army's Forecast and Allocation of Army Recruiting Resources (FAARR) model. Using sensitivity analysis, validation forecasting, and Monte-Carlo simulation of a known production function, this research demonstrated the FAARR model in its current form does not provide accurate forecasts of GSMA contract production.

The FAARR model can not be used for "what if" analysis and is not accurate when recruiting resources change significantly from current levels. The FAARR model uses the estimated values from a *descriptive, non-parametric* DEA model as production function parameters in a *prescriptive* forecasting model. The FAARR model's assumptions, such as the restrictions placed on the relative value of the recruiting resources, are invalid. Also, the model's 2nd stage mathematical programming formulation is not an optimization model as indicated in its documentation (13:10).

The FAARR model's GSMA contract forecasts were ultra-sensitive to the recruiting resource levels used in the actual forecast and to the specification of the first phase DEA model. A relatively small 5% increase in the aggregate level of all recruiting resources resulted in a 42% increase in forecasted GSMA production. Analysis and experience indicate this is a grossly unrealistic increase in forecasted production given the minimal increase in recruiting resources. Additionally, without any constraints on the DEA virtual multipliers, the FAARR estimated multipliers produced a model unable to find a feasible solution for the production forecast.

The FAARR model was also not able to accurately forecast actual contract production using historical production data. The FAARR model's best validation forecast had a MAPE of 14% with an average recruiting battalion MAPE of 46% and a maximum battalion MAPE of 153%. Individual recruiting battalion forecasts had extremely large errors. Simple time-series models provided more accurate forecast estimates than the FAARR model. In fact, the Naive Forecast 1 model actually provided more accurate production forecasts than the FAARR model for all Measures Of Effectiveness (MOE).

Finally, using a simulation of a known CRS production function, the FAARR model was not able to accurately classify simulated battalions as efficient or inefficient or accurately estimate the actual battalion efficiency scores. The FAARR model incorrectly classified efficient battalions as inefficient 95% of the time. Average FAARR model estimated battalion efficiency scores for two different production functions were 0.54 and 0.64 when the actual average simulated battalion efficiency was 0.88. Additionally, the average correlation coefficient between each battalion's actual and estimated efficiencies for the two models was only 0.13 and 0.16, respectively.

The FAARR model assumes a VRS production function underlies the Army recruiting process. This research indicates that if the actual production process is not VRS, FAARR model forecasts will not be accurate. The FAARR model incorrectly attributed simulated battalions' less than efficient production output to a change in the production process's returns-to-scale and not to actual battalion inefficiency.

The second part of this research developed a three phase strategy to select the most accurate DEA model formulation for the Army recruiting process. Using this strategy, a

five variable CRS CCR model was identified as the most accurate DEA model to estimate U.S. Army recruiting battalion efficiency. This model provided significantly different results than the current FAARR eight variable, Multiplicative VRS DEA model.

The DEA model building strategy which was developed used multi-variate statistical analysis and OLS regression to select relevant input variables. A Monte-Carlo simulation of a production function using these relevant input variables was then used to select the most appropriate shape of the DEA envelopment frontier. This research illustrated how multi-variate statistical techniques can be combined with expert opinion to make decisions on whether or not to include specific input variables in a DEA model

This research's results concerning the selection of the shape of the DEA envelopment frontier are similar to other simulation studies of DEA model misspecification. Using four simulated IRS production processes, the CCR model was the most accurate DEA envelopment. The CCR model incorrectly classified 24% of all simulated battalions as efficient or inefficient. In contrast, the BCC model overestimated battalion efficiency scores by 6.25% and incorrectly classified 42% of all battalions. An incorrect choice of a VRS DEA model for an IRS production process resulted in upwardly biased efficiency estimates. In an attempt to maximize each battalion's efficiency score, the VRS models attribute a battalion's less than efficient output production to a change in the production process's returns-to-scale and not to inherent battalion inefficiency. This research demonstrated the choice of a particular DEA model implies an assumption about the production processes' returns-to-scale properties and is critical in accurately estimating DMU efficiency.

This research assumed the recruiting process's incentive structure and recruiter behavior is such that all recruiters would seek to maximize GSMA contract production given any allocation of recruiting resources. This assumption may not be valid because of the process USAREC uses to assign specific recruiting battalion production missions and the way recruiters react to these production missions. However, DEA models may still be used to determine the relative efficiency of recruiting battalions even if recruiters do not attempt to maximize their contract production.

In conclusion, this research summarizes much of the theory and current practice of DEA modeling and provides the Operations Research community an appropriate strategy to build accurate DEA models.

5.2 Improving USAREC Econometric and Forecasting Models

The results of this research indicate a five variable CCR DEA model using recruiters, population, unemployment, print GRPs, and local advertising expenditures may be the most accurate model to evaluate Army recruiting battalion efficiency. Accurate recruiting battalion efficiency information from the CCR DEA model can be used in multiple stage mathematical or statistical models. DEA provides an additional variable to be used with any current or future USAREC forecasting model to improve forecast accuracy or improve resource parameter estimates.

Additionally, this research has also demonstrated that in the short term, simple, time-series forecasting models may be more accurate than econometric based causal models. However, time-series models can not be used to estimate resource elasticity parameters. USAREC forecasts and parameter estimates may be improved by developing and using

two totally separate models. A simple time-series model may be used for contract forecasts and a more complex econometric model may be used for resource elasticity parameter estimates.

5.3 Extensions of Current Research

This research may be expanded in a number of different directions. First, this research may be expanded by using additional quarterly recruiting data as the data becomes available. This research evaluated the various forecasting models over three consecutive quarters. New data will provide more information regarding the robustness and accuracy of the various forecasting model estimates over a wider time frame.

Another direction may involve changing the specific simulation used to select the most accurate DEA envelopment. Future experimental designs using Response Surface Methodology (RSM) techniques may be developed to estimate the sensitivity of the accuracy of the DEA efficiency estimates for varying probability distributions of the efficiency scores for the simulated DMUs. The current research used a truncated, normal distribution for the efficiency scores of the simulated DMUs estimated from the FAARR DEA model. Future research may determine if the identified, most accurate DEA envelopment is sensitive to the efficiency score probability distribution used in the simulation. Additionally, RSM experimental designs may also be used to determine the sensitivity of the accuracy of the DEA estimates to changes in the number of input variables or to more complex types of production functions such as stochastic frontier functions.

Future researchers may also decide to use the Spearman rank correlation coefficient instead of the ordinary correlation coefficient as a MOE to evaluate the accuracy of various DEA model formulations. Viewed in the context of ranking and selection theory, the Spearman rank correlation coefficient may be a more appropriate measure of the ability of a specific envelopment to estimate a DMU's true efficiency. However, discriminating between efficient DMUs-- all with an efficiency score of 1-- and assigning the DMUs the appropriate rank may be difficult.

A third direction for future research may include the evaluation of forecast accuracy using DEA efficiency information in more complex and detailed forecasting models--either current USAREC models or models to be developed in the future. Using a more complex forecasting model may result in a more significant test of the hypothesis that including DEA efficiency information improves model forecasts and parameter estimates.

Finally, another area for possible future research is to compare recruiting battalion efficiency estimates from CRS and VRS models-- comparing the CCR model to the BCC model, or the Multiplicative model to the Multiplicative without intercept model. These different comparisons can be used to positively identify subsets of efficient and inefficient battalions regardless of the production processes' returns-to-scale classification. This research illustrated that invalid assumptions concerning a production processes' returns-to-scale may result in inaccurate efficiency estimates. Given the different returns-to-scale assumptions for the CRS and VRS models, the CRS model identifies a conservative, lower limit of the number of efficient DMUs. In contrast, a VRS model identifies a conservative, lower limit of inefficient DMUs. By using these two type of models simultaneously,

analysts can positively identify a subset of all evaluated DMUs as either efficient-- using a CRS DEA model-- or inefficient-- using a VRS DEA model-- even if the production processes' return-to-scale property can not be accurately identified. Therefore, the analyst does not need to make or prove any assumption concerning the production processes' returns-to-scale, and can still positively identify some DMUs as efficient or inefficient. If management intends to use DEA efficiency analysis qualitatively to identify best and worst operating practices, this positively identified subset of efficient and inefficient DMUs may provide adequate information.

Appendix A. Sample GAMS Simulation Model

```

$title Combined Optimization/Simulation Model
$ontext
    New model -- OCT 1997
    Coded by   : Piskator, Gene M. CPT (from original DEA model by Yuying Wang)
$offtext
$offsymxref offsymlist offuellist offuelxref

option solprint = off;
option decimals = 4;
option limrow = 0, limcol = 0;
option seed = 235357;

SETS DMU / 1A, 1B, 1D , 1E, 1G, 1K, 1L, 1N, 1O, 3A, 3D, 3E , 3G,
3H, 3I, 3J, 3N, 3T, 4C, 4E, 4G, 4I, 4J, 4K, 4L, 4N, 5A, 5C, 5D, 5H,
5I, 5J, 5K , 6D, 6F, 6G, 6H, 6I, 6J, 6K, 6L /
*assigns DMU names/identifies DMUs

DATANAMES / OPR, MAGZINE, local, pop, unemp, GSMA /
*assigns variable names

IN (DATANAMES) / OPR, magazine, local, pop, unemp /
*names input variables

OUT (DATANAMES) / GSMA /;
*names output variable

SET ITER /1*100/;
*sets number of Monte-Carlo simulation replications

Alias (in, I);
Alias (i,j);
Alias (in, KK);
Alias (out, R);
Alias (DMU, DMUcurr);
*assigns alias for data sets

TABLE COVAR (I,J)
      OPR  MAGZINE  LOCAL  POP  UNEMP
opr    25.556    0      0      0      0
magzine -5.957   59.833    0      0      0
local   3.27E+03  852.028  5.89E+03    0      0
pop     1.43E+04 -4.38E+03  9.62E+03  2.70E+04    0

```

unemp 0.271 0.094 -0.17 0.195 1.059

;

*covariance table for input variables. Used with Cholesky decomposition to generate

*multi-variate normal vector of input variables. See Law, Averill M. and W. David

*Kelton. Simulation Modeling and Analysis. New York: McGraw-Hill, 1991, pages 505-

*506.

PARAMETERS

X0(I) inputs of DMU under evaluation

Y0(R) outputs of DMU under evaluation

X(DMU,I)

Y(DMU,R)

NORMRAN(I)

MVNORM(DMU,I)

coefuu(DMU, r) table of uu (output virtual multipliers)

coefvv(DMU, i) table of vv (input virtual multipliers)

obj(DMU) table of objval (efficiency scores)

RANEFFIC(DMU)

AVEFFIC

DEVEFFIC(DMU)

AVEDEV

TOTGSMA

AVEOPR

AVETVGRP

AVERDGRP

AVEPGRP

AVELOCL

AVEDOD

AVEPOP

AVEUNEMP

AVEU0

KNOWNEFF

CORRCOEF

NUMER

DENOM

DENOM1

DENOM2

NUMEFF

IDEFF

TRUEFF(DMU)

ESTEFF(DMU)

COUNT1

COUNT2

COUNT3

COUNT4

AVEINPUT(I) / OPR 126.2926829
MAGZINE 250.9756098
LOCAL 15978.12195
POP 100331.5366
UNEMP 4.73195122

/;

*mean input level for resource variables used to generate Multi-variate normal resource

*vectors assigned to simulated, random DMUs (recruiting battalions)

VARIABLES

OBJVAL objective values

OUTPUT;

POSITIVE VARIABLES

UUY(r)

VVX(i)

CONTRACTS(DMU);

equations

OBJFCN3

const3(dmu)

lessonein(i)

lessoneout(r)

NORM

;

OBJFCN3.. objval =e= sum(i, x0(i)*vvx(i));

*CCR DEA objective function maximizes value of outputs

const3(DMU).. SUM(r, -uuy(r)*y(DMU,r)) + SUM(i, vvz(i)*x(DMU,i)) =g= 0;

*CCR DEA constraint

lessonein(i) .. vvz(i) =g= .000001;

lessoneout(r) .. uuy(r) =g= .000001;

*Non-archimedean infinitesimal constraint

norm .. sum(r, y0(r)*uuy(r)) =e= 1;

*output variable normalization constraint

MODEL DEA3 /OBJFCN3, const3, lessonein, lessoneout, NORM/;

file DEASIM3; put DEASIM3; DEASIM3.pc=5; DEASIM3.nd=4;

*names output file

loop (i, put i.tl);

put 'U0', 'GSMA's', 'EST-EFF', 'KNOWN-EFF', 'EFF-ERROR', 'CORR-COEFF',
'EFF|EFF', 'N-EFF|N-EFF', 'N-EFF|EFF', 'EFF|N-EFF';

*writes simulation statistics information to output file

LOOP (ITER,

COUNT1=0;

COUNT2=0;

COUNT3=0;

COUNT4=0;

*reset all DMU identification counters to zero. These counters are used to classify

*DMUs as truly efficient or inefficient and estimated as efficient or inefficient

LOOP(DMU,

TRUEFF(DMU) =0;

ESTEFF(DMU)=0;

NORMRAN(I)=NORMAL(0,1);

X(DMU,I) = (AVEINPUT(I) + SUM(J, (NORMRAN(J)*COVAR(I,J))));

*generates multi-variate normal input vector using Chosleky decomposition

LOOP(I,

IF (X(DMU,I) LT 2.5, X(DMU,I)=2.5);

);

RANEFFIC(DMU)=(NORMAL(.8875,.0923));

IF (RANEFFIC(DMU) GT 1, RANEFFIC(DMU)=1.0);

IF(RANEFFIC(DMU) = 1, TRUEFF(DMU)=1);

IF (RANEFFIC(DMU) LT 0, RANEFFIC(DMU)=0.01);

);

*randomly generates true efficiency scores from truncated normal distribution

X(DMU,I) = LOG(X(DMU,I));

RANEFFIC(DMU)=LOG(RANEFFIC(DMU));

Y(DMU,"GSMA")=(RANEFFIC(DMU) -1.842005608 +
.991010447*X(DMU,"OPR") + 0.025509477*X(DMU,"local") +
.176151048*X(DMU,"unemp") + 0.119530613*X(DMU,"magzine") +
0.096590765*X(DMU,"pop"));

X(DMU,I) = EXP(X(DMU,I));

Y(DMU,"GSMA")= EXP(Y(DMU,"GSMA"))+normal(0,10.00);

*calculates GSMA production for each DMU based on production function, random input

*vector, true efficiency score, and normally distributed error term

```

LOOP(DMUcurr,
  x0(i)=x(dmucurr,i);
  y0(r)=y(dmucurr,r);

  SOLVE DEA3 USING LP Minimizing OBJval;

  coefuu(DMUcurr, R) = uuy.L(R);
  coefvv(DMUcurr, I) = vvx.L(I);
  obj(DMUcurr) = (1/objval.L);
  IF(OBJ(DMUCURR) GT .9999, ESTEFF(DMUCURR)=1);
);
*solves CCR DEA model for each DMU

LOOP(DMU,
  IF ((TRUEFF(DMU) = 1) AND (ESTEFF(DMU) =1), COUNT1=COUNT1+1);
  IF ((TRUEFF(DMU) = 0) AND (ESTEFF(DMU) =0), COUNT2=COUNT2+1);
  IF ((TRUEFF(DMU) = 1) AND (ESTEFF(DMU) =0), COUNT3=COUNT3+1);
  IF ((TRUEFF(DMU) = 0) AND (ESTEFF(DMU) =1), COUNT4=COUNT4+1);
);
*counts DMUs to identify which were correctly classified by the DEA model

AVEFFIC=( (SUM(DMU,OBJ(DMU))) /41.0) ;
DEVEFFIC(DMU)=ABS((OBJ(DMU))-EXP(RANEFFIC(DMU)));
KNOWNEFF=(SUM(DMU, EXP(RANEFFIC(DMU)))/41.0);
AVEDEV=(SUM(DMU, DEVEFFIC(DMU)))/41.0;
TOTGSMA=SUM(DMU, (Y(DMU, "GSMA")));
AVEOPR=(SUM(DMU, coefVV(DMU, "OPR"))/41.0);
AVElocl=(SUM(DMU, coefVV(DMU, "local"))/41.0);
AVEPGRP=(SUM(DMU, coefVV(DMU, "MAGZINE"))/41.0);
AVEunemp=(SUM(DMU, coefVV(DMU, "unemp"))/41.0);
AVEPOP=(SUM(DMU, coefVV(DMU, "POP"))/41.0);
NUMER= SUM(DMU, ((OBJ(DMU))-AVEFFIC) * (EXP(RANEFFIC(DMU))-KNOWNEFF) );
DENOM1=SUM(DMU,(( (OBJ(DMU))-AVEFFIC)*((OBJ(DMU))-AVEFFIC) ));
DENOM2= SUM(DMU,(( EXP(RANEFFIC(DMU))-KNOWNEFF)*(EXP(RANEFFIC(DMU))-KNOWNEFF) ));
DENOM=SQRT(DENOM1*DENOM2);
CORRCOEF= NUMER/DENOM;
*calculates average efficiency scores and correlation coefficients between true and estimated efficiency scores

put /;
PUT AVEOPR:10:4;
PUT AVEPGRP:10:4;

```

PUT AVElocl:10:4;
PUT AVEPOP:10:4;
PUT AVEunemp:10:4;
PUT 'no intcpt';
PUT TOTGSMA:10:4;
PUT AVEFFIC:10:4;
PUT KNOWNEFF:10:4;
PUT AVEDEV:10:4;
PUT CORRCOEF:10:4;
PUT COUNT1:10:4;
PUT COUNT2:10:4;
PUT COUNT3:10:4;
PUT COUNT4:10:4;
*outputs MOEs and average scores to file
);

Bibliography

1. Andersen, Per and Niels Christian Petersen. "A Procedure for Ranking Efficient Units in Data Envelopment Analysis," Management Science, 39(10): 1261-1264 (October 1993).
2. Arnold, Victor L., Indranil Bardhan, William Cooper, and Subal C. Kumbhakar. "New Uses of DEA and Statistical Regressions for Efficiency Evaluation and Estimation--with an Illustrative Example to Public Secondary Schools in Texas," Annals of Operations Research, Vol 66: 255-277 (1996).
3. "Attrition Big Problem", The Army Times, 11 March 1997:22.
4. Banker, Rajiv D. "Maximum Likelihood, Consistency and Data Envelopment Analysis: A Statistical Foundation", Management Science, 39(10): 1265-1273 (October 1993).
5. Banker, Rajiv D., Hsihui Chang, and W.W. Cooper. "Simulation Studies of Efficiency, Returns to Scale and Misspecification with Nonlinear Functions in DEA", Annals of Operations Research, 66:233-253 (1996).
6. Banker, Rajiv D., Robert F. Conrad, and Robert P. Strauss. "A Comparative Application of Data Envelopment Analysis and Translog Methods: An Illustrative Study of Hospital Production", Journal of Management Science, 25:30-43 (1986).
7. Banker, Rajiv D. and Richard C. Morey. "Incorporating Value Judgments in Efficiency Analysis," Research in Governmental and Nonprofit Accounting, 5: 245-267 (1989).
8. Bardhan, Indranil R., William W. Cooper, and Subal C. Kumbhakar. "A Simulation Study of Joint Uses of Data Envelopment Analysis and Statistical Regressions for Production Function Estimation and Efficiency Evaluation", to be published in The Journal of Productivity Analysis (1998).
9. Barr, Richard S., Lawrence M. Seiford, and Thomas F. Siems. "An Envelopment-Analysis Approach to Measuring the Managerial Efficiency of Banks", Annals of Operations Research, 45: 1-19 (1993).
10. Bessent A., W. Bessent, J. Kennington, and B. Reagan. "An Application of Mathematical Programming to Assess Productivity in the Houston Independent School District", Management Science, 28 (12): 1355-1367 (December 1982).
11. Boussofiane, A., R.G. Dyson and E. Thanassoulis. "Applied Data Envelopment Analysis," European Journal of Operational Research, 52: 1-15 (1991).

12. Bowlin, W.F., A. Charnes, W.W. Cooper, and H. D. Sherman. "Data Envelopment Analysis and Regression Approaches to Efficiency Estimation and Evaluation", Annals of Operations Research, 2: 112-138 (1985).
13. Brockett, P.L., J.J. Rousseau, Y. Wang, and G. Yu. The FAARR Decision Support System. Austin TX: Center for Cybernetic Studies, January 1997.
14. Brooke, Anthony, David Kendrick, and Alexander Meerhaus. GAMS: A Users Guide. Danvers, MA: The Scientific Press, 1992.
15. Callen, Jeffrey L. "Data Envelopment Analysis: Partial Survey and Applications for Management Accounting", Journal of Management Accounting Research, Fall 1991.
16. Charnes, Abraham, W.W. Cooper, Arie Y. Lewin and Lawrence M. Seiford. Data Envelopment Analysis: Theory, Methodology and Applications. Boston: Kluwer Academic Publishers, 1994.
17. Charnes, Abraham, W.W. Cooper, A.Y. Lewin, R.C. Morey and J. Rousseau. "Sensitivity and Stability Analysis in DEA," Annals of Operations Research, 2: 139-156 (1985) (AD-A159-586).
18. Charnes, Abraham, W. W. Cooper and A. P. Schinnar. Theorem on Homogeneous Functions and Extended Cobb-Douglas Forms. Proceedings of the National Academy of Science, 73(10): 3747-3748 (October 1976).
19. Charnes, Abraham, W.W. Cooper, M. Sears and S. Zlobec. "Efficiency Evaluations in Perturbed Data Envelopment Analysis," Proceedings of the Second International Conference on Parametric Optimization and Related Topics, Mathematical Research: 62:39-50. June 27-July 2 1989.
20. Charnes, Abraham, Gallegos, Armando, and Li, Hongyu, "Robustly Efficient Parametric Frontiers via Multiplicative DEA for Domestic and International Operations of the Latin American Airline Industry," European Journal of Operational Research, Vol. 88, (1996). pp. 525-536.
21. Charnes, Abraham, Z.M. Huang, J. Semple, T. Song and D. Thomas. "Origins and Research in Data Envelopment Analysis," The Arabian Journal of Science and Engineering, 15 (4B): 617-625 (October 1990).
22. Department of Defense. "Military Personnel Policy Major Initiatives." n. pag. WWWeb, <http://dticaw.dtic.mil/prhome/mpp-page.html>. 20 September 1997.
23. Dillon, William R. and Matthew Goldstein. Multivariate Analysis: Methods and Applications. New York: John Wiley and Sons, Inc., 1984.

24. Donohue, Joan M., Ernest C. Houck and Raymond Myers. "Sequential Experimental Designs for Simulation Metamodeling". Proceedings of the 1992 Winter Simulation Conference. 539-547.
25. Dornbusch, Rudiger, and Stanley Fischer. Macroeconomics. New York: McGraw Hill, Inc., 1987.
26. Farrell, M.J. "The Measurement of Productive Efficiency," Journal of the Royal Statistical Society, 120(3): 253-290 (1957).
27. Golany, Boaz, Phillips, F. Y., and Rousseau, J. J., "Models for Improved Effectiveness Based on DEA Efficiency Results," IIE Transactions, Vol. 25, No. 6, (November 1993). pp. 2-10.
28. Golany, Boaz and Eran Tamir. "Evaluating Efficiency-Effectiveness-Equality Trade-offs: A Data Envelopment Analysis Approach," Management Science, 41 (7): 1172-1184 (July 1995).
29. Golany, Boaz, and Yu, Gang, "A Goal Programming-Discriminant Function Approach to the Estimation of an Empirical Production Function Based on DEA Results," The Journal of Productivity Analysis, Vol. 6, No. 2, (1995). pp. 171-186.
30. Higle, Julia L., Wing W. Lowe and Ronald Odio. "Conditional Stochastic Decomposition: An Algorithmic Interface for Optimization and Simulation," Operations Research, 42(2): 311-322 (March-April 1994).
31. Horsky, Dan and Paul Nelson. "Evaluation of Salesforce Size and Productivity Through Efficient Frontier Benchmarking," Marketing Science, 15(4): 301-320 (1996).
32. Jaska, Patrick V. Sensitivity for Data Envelopment Analysis with an Application to Texas Agricultural Production. Ph.D. dissertation. University of Texas at Arlington, Arlington TX, December 1991.
33. Johnston, J. Econometric Methods, New York: McGraw-Hill Co., 1984.
34. Kachigan, Sam Kash. Multivariate Statistical Analysis, New York: Radius Press Co., 1991.
35. Kennedy, Peter. A Guide to Econometrics. Cambridge, Mass.: MIT Press, 1985.
36. Law, Averill M. and W. David Kelton. Simulation Modeling and Analysis. New York: McGraw-Hill, 1991.

37. Makridakis, Spyros, Steven C. Wheelwright, and Victor McGee. Forecasting: Methods and Applications. New York: John Wiley and Sons, Inc., 1983.
38. Maurice, S. Charles, Owen R. Phillips, and C. E. Ferguson. Economic Analysis: Theory and Application, Homewood, IL: Richard D. Irwin, Inc., 1982.
39. Maze, Rick. "Services Want More Money for Recruiting Efforts," Army Times, 15 April 15, 1996: 16.
40. Myers, Raymond H., and Douglas C. Montgomery. Response Surface Methodology. New York: John Wiley and Sons, Inc., 1995.
41. Nielsen Media Research. National Reference Supplement 1995-1996. Nielsen Media Research, Inc., 1995.
42. Olesen, O.B., and N. C. Petersen, N.C. "Chance Constrained Efficiency Evaluation", Management Science, 41(3): 442-457 (March 1995).
43. Sengupta, Jati K. "Tests of Efficiency in Data Envelopment Analysis", Computer Operations Research, 17(2): 123-132, 1990.
44. Sengupta, Jati K. "Recent Models in Data Envelopment Analysis: Theory and Applications". Applied Stochastic Models and Data Analysis, 1-26, New York: John Wiley and Sons, LTD, 1996.
45. Simons, John. "Unemployment Rate Fell in November," The Wall Street Journal, 8 December 1997, sec. A2.
46. Smith, Peter. "Model Misspecification in Data Envelopment Analysis," Annals of Operations Research, Vol 66: 233-252 (1997).
47. Stolp, Chandler. "Strength and Weaknesses of Data Envelopment Analysis: An Urban and Regional Perspective", Computing, Environmental and Urban Systems, 14: 103-116, 1990.
48. Tice, Jim. "New Initiatives in Recruiting", The Army Times, 17 March 1997: 15.
49. Tice, Jim. "Inbrief", The Army Times, 7 July 1997: 22.
50. Thomas, David A. Data Envelopment Analysis Methods in the Management of Personnel Recruitment Under Competition in the Context of U.S. Army Recruiting. Ph.D. dissertation. University of Texas at Austin, Austin TX, August 1990.

51. United States Army Recruiting Command. U.S. Army Recruiting Command Overview, 21 April 1997.
52. Valdmanis, Vivian. "Sensitivity Analysis for DEA Models: An Empirical Example Using Public vs NFP Hospitals," Journal of Public Economics, 48: 185-205 (July 1992).
53. Weigley, Russel F. History of the United States Army. Bloomington, IN: Indiana University Press, 1984.
54. Zimmerman, Hans-Jurgen, and Michael D. Sovereign. Quantitative Models for Production Management. Englewood Cliffs, NJ: Prentice-Hall Inc., 1974.

Vita

Captain Gene M. Piskator was born on November 20, 1964 in Worcester, Massachusetts. He enlisted in the Army in 1982 and graduated from the United States Military Academy at West Point, NY, in 1988 with a Bachelor of Science degree in Economics. Commissioned in the Infantry, he served as an airborne infantry rifle platoon leader, anti-armor platoon leader, and rifle company executive officer. Following his branch transfer to the Transportation Corps, he served as a transportation battalion adjutant, a heavy boat company commander, and a personnel staff officer in the Office of the Chief of Transportation. He entered the Graduate School of Engineering, Air Force Institute of Technology in August of 1996. He is married to the former Tara J. Mullins.

Permanent address: 25 Toryfort Lane
Worcester, MA 01602

E-mail address: gene1tara@aol.com

REPORT DOCUMENTATION PAGE			Form Approved OMB No. 0704-0188	
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.				
1. AGENCY USE ONLY (Leave blank)		2. REPORT DATE March 1988	3. REPORT TYPE AND DATES COVERED Masters Thesis	
4. TITLE AND SUBTITLE Verification and Validation of FAARR Model and Data Envelopment Analysis Models for United Sates Army Recruiting			5. FUNDING NUMBERS	
6. AUTHOR(S) Gene M. Piskator Captain, U.S. Army				
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Graduate School of Engineering Air Force Institute of Technology Wright-Patterson AFB, Ohio 45433-6583			8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) United States Army Recruiting Command Commander, USAREC ATTN: PAE 1307 3rd Avenue Fort Knox, KY 40121			10. SPONSORING/MONITORING AGENCY REPORT NUMBER	
11. SUPPLEMENTARY NOTES Advisor: LTC Jack M. Kloeber phone: Commercial (937) 255-6565 jkloeber@afit.af.mil				
12a. DISTRIBUTION AVAILABILITY STATEMENT Approved for public release; distribution unlimited			12b. DISTRIBUTION CODE	
13. ABSTRACT (Maximum 200 words) This research has two objectives—to verify and validate the U.S. Army's Forecast and Allocation of Army Recruiting Resources (FAARR) model and to develop a Data Envelopment Analysis (DEA) modeling strategy. First, the FAARR model was verified using a simulation of a known production function and validated using sensitivity analysis and ex-post forecasts. FAARR model forecasts were not accurate and were extremely sensitive to any changes in the model's linear programming constraints and to changes in recruiting resource levels. Second, this research describes a three phase modeling strategy to build accurate DEA models. DEA has become a popular tool to evaluate the relative efficiency of many types of organizations. However, the literature has paid little attention to the practical problems of selecting the appropriate input variables and envelopment frontier. Analysts may use a number of diagnostic techniques to detect misspecification in statistics based models. No such diagnostics exist for DEA models. Without a-priori knowledge concerning the production process's appropriate input variables and returns to scale, analysts do not know if they have constructed an accurate DEA model. Using a three-phase strategy, relevant DEA model input variables are selected using Principal Component Analysis and Ordinary Least Squares (OLS) regression. The appropriate DEA envelopment frontier is selected using a Monte-Carlo simulation of an estimated production function representing the actual production process. The research concludes by demonstrating ex-post forecasts from a combined OLS/DEA model were more accurate when the DEA model selected by the three phase modeling strategy was used.				
14. SUBJECT TERMS Data Envelopment Analysis (DEA), Army recruiting, production functions			15. NUMBER OF PAGES 120	
			16. PRICE CODE	
17. SECURITY CLASSIFICATION OF REPORT Unclassified	18. SECURITY CLASSIFICATION OF THIS PAGE Unclassified	19. SECURITY CLASSIFICATION OF ABSTRACT Unclassified	20. LIMITATION OF ABSTRACT UL	

GENERAL INSTRUCTIONS FOR COMPLETING SF 298

The Report Documentation Page (RDP) is used in announcing and cataloging reports. It is important that this information be consistent with the rest of the report, particularly the cover and title page. Instructions for filling in each block of the form follow. It is important to *stay within the lines* to meet *optical scanning requirements*.

Block 1. Agency Use Only (Leave blank).

Block 2. Report Date. Full publication date including day, month, and year, if available (e.g. 1 Jan 88). Must cite at least the year.

Block 3. Type of Report and Dates Covered. State whether report is interim, final, etc. If applicable, enter inclusive report dates (e.g. 10 Jun 87 - 30 Jun 88).

Block 4. Title and Subtitle. A title is taken from the part of the report that provides the most meaningful and complete information. When a report is prepared in more than one volume, repeat the primary title, add volume number, and include subtitle for the specific volume. On classified documents enter the title classification in parentheses.

Block 5. Funding Numbers. To include contract and grant numbers; may include program element number(s), project number(s), task number(s), and work unit number(s). Use the following labels:

C - Contract	PR - Project
G - Grant	TA - Task
PE - Program Element	WU - Work Unit Accession No.

Block 6. Author(s). Name(s) of person(s) responsible for writing the report, performing the research, or credited with the content of the report. If editor or compiler, this should follow the name(s).

Block 7. Performing Organization Name(s) and Address(es). Self-explanatory.

Block 8. Performing Organization Report Number. Enter the unique alphanumeric report number(s) assigned by the organization performing the report.

Block 9. Sponsoring/Monitoring Agency Name(s) and Address(es). Self-explanatory.

Block 10. Sponsoring/Monitoring Agency Report Number. (If known)

Block 11. Supplementary Notes. Enter information not included elsewhere such as: Prepared in cooperation with....; Trans. of....; To be published in.... When a report is revised, include a statement whether the new report supersedes or supplements the older report.

Block 12a. Distribution/Availability Statement.

Denotes public availability or limitations. Cite any availability to the public. Enter additional limitations or special markings in all capitals (e.g. NOFORN, REL, ITAR).

DOD - See DoDD 5230.24, "Distribution Statements on Technical Documents."

DOE - See authorities.

NASA - See Handbook NHB 2200.2.

NTIS - Leave blank.

Block 12b. Distribution Code.

DOD - Leave blank.

DOE - Enter DOE distribution categories from the Standard Distribution for Unclassified Scientific and Technical Reports.

NASA - Leave blank.

NTIS - Leave blank.

Block 13. Abstract. Include a brief (*Maximum 200 words*) factual summary of the most significant information contained in the report.

Block 14. Subject Terms. Keywords or phrases identifying major subjects in the report.

Block 15. Number of Pages. Enter the total number of pages.

Block 16. Price Code. Enter appropriate price code (*NTIS only*).

Blocks 17. - 19. Security Classifications. Self-explanatory. Enter U.S. Security Classification in accordance with U.S. Security Regulations (i.e., UNCLASSIFIED). If form contains classified information, stamp classification on the top and bottom of the page.

Block 20. Limitation of Abstract. This block must be completed to assign a limitation to the abstract. Enter either UL (unlimited) or SAR (same as report). An entry in this block is necessary if the abstract is to be limited. If blank, the abstract is assumed to be unlimited.